electronics

MDPI

Article

Using the Buckingham π Theorem for Multi-System Transfer Learning: A Case-Study with 3 Vehicles Sharing a Database

William Therrien, Olivier Lecompte and Alexandre Girard *

Département de Génie Mécanique, Faculté de Génie, Université de Sherbrooke, 2500, boul. de l'Université, Sherbrooke, QC J1K 2R1, Canada; william.therrien@usherbrooke.ca (W.T.); olivier.lecompte@usherbrooke.ca (O.L.)
* Correspondence: alexandre.girard2@usherbrooke.ca

Abstract: Many advanced driver assistance schemes or autonomous vehicle controllers are based on a motion model of the vehicle behavior, i.e., a function predicting how the vehicle will react to a given control input. Data-driven models, based on experimental or simulated data, are very useful, especially for vehicles difficult to model analytically, for instance, ground vehicles for which the ground-tire interaction is hard to model from first principles. However, learning schemes are limited by the difficulty of collecting large amounts of experimental data or having to rely on high-fidelity simulations. This paper explores the potential of an approach that uses dimensionless numbers based on Buckingham's π theorem to improve the efficiency of data for learning models, with the goal of facilitating knowledge sharing between similar systems. A case study using car-like vehicles compares traditional and dimensionless models on simulated and experimental data to validate the benefits of the new dimensionless learning approach. Preliminary results from the case study presented show that this new dimensionless approach could accelerate the learning rate and improve the accuracy of the model prediction when transferring the learned model between various similar vehicles. Prediction accuracy improvements with the dimensionless scheme when using a shared database, that is, predicting the motion of a vehicle based on data from various different vehicles was found to be 480% more accurate for predicting a simple no-slip maneuver based on simulated data and 11% more accurate to predict a highly dynamic braking maneuver based on experimental data. A modified physics-informed learning scheme with hand-crafted dimensionless features was also shown to increase the improvement to precision gains of 917% and 28% respectively. A comparative study also shows that using Buckingham's π theorem is a much more effective preprocessing step for this task than principal component analysis (PCA) or simply normalizing the data. These results show that the use of dimensionless variables is a promising tool to help in the task of learning a more generalizable motion model for vehicles, and hence potentially taking advantage of the data generated by fleets of vehicles on the road even though they are not identical.

Keywords: dimensional analysis; transfer learning; machine learning; robotics; dynamics

check for updates

Citation: Therrien, W.; Lecompte, O.; Girard, A. Using the Buckingham π Theorem for Multi-System Transfer Learning: A Case-Study with 3 Vehicles Sharing a Database. *Electronics* **2024**, *13*, 2041. https://doi.org/10.3390/electronics13112041

Academic Editors: Nikolay Hinov, Jožef Ritonja and Darius Andriukaitis

Received: 18 April 2024
Revised: 14 May 2024
Accepted: 21 May 2024
Published: 23 May 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The typical approach to solving motion control problems for autonomous vehicles and other types of robotic systems is to use a physics-based kinematic or dynamic model of the system to plan trajectories and establish feedback laws [1–3]. With the rise of efficient machine learning algorithms, much effort has been spent in using learning in planners and controllers so that autonomous vehicles can learn from experience and improve over time [4–6]. However, major difficulties limit the success of the learning scheme when it comes to controlling real physical platforms. One of the major challenges is the difficulty of collecting the large amount of experimental data required for advanced learning schemes such as deep-learning [6,7]. A popular solution is to generate data using high-fidelity simulations instead of experiments. Nevertheless, this approach also has

multiple challenges, such as the difficulty of capturing certain behaviors in a simulator and transferring the results to the real world [8]. All in all, data efficiency is a critical bottleneck.

Sharing learned data and policies across multiple systems could help solve the problem of collecting huge amounts of data. For example, instead of having to start learning from scratch, a centralized database for a fleet of vehicles would enable the newly deployed vehicle to use knowledge based on thousands of hours of motion data collected by hundreds of vehicles. However, this is difficult to achieve when the studied systems do not share identical characteristics and, therefore, the appropriate control policy differs. This paper investigates the potential of leveraging dimensionless numbers and the Buckingham π theorem to help learn more general models. Section 2 presents some related work and discusses the original contribution of this paper. Section 3 defines the problem used as the case study in this paper. Section 4 presents the learning results based on simulated data and Section 5 presents the learning results based on an experimental validation with three small-scale vehicles shown in Figure 1.



(**a**) Small vehicle (**b**) Long vehicle (**c**) Large vehicle

Figure 1. Small scale platforms of various dimensions used for experimental validation.

2. Background

Multi-task transfer learning has recently received significant research attention [9–11]. This paper focuses on multi-system transfer where the objective is to share knowledge between similar systems that accomplish the same task. Specifically for multi-system transfer, an interesting approach is presented in [12] where a shared database can outperform a robot-specific training approach that uses 4×–20× more data. In [13], the authors address the problem of inappropriate transfer, which can result in a decrease in performance. By assessing how similar the two systems are, it is possible to know with which systems the transfer learning would be most beneficial. The closest related work is probably [14], where the authors tackle the problem directly by guiding the policy network with a representation vector built from the information on the hardware. The goal is similar in this work, but the approach is different: instead of two vectors, one containing the states and the other containing the characteristics of the hardware, dimensional analysis is used to build a single dimensionless vector using the Buckingham π theorem.

Dimensional analysis [15–17], i.e., the Buckingham π theorem is a useful tool to simplify input–output relationships that involve physical quantities. It is used extensively in the field of fluid mechanics to minimize the number of experiments required to characterize a behavior. There is a recently renewed interest in using dimensional analysis in the context of learning. It was proposed to pre-process data with the Buckingham π theorem to improve the robustness and accuracy of learned complex turbulent flow dynamics [18,19]. Also, automated methods were proposed to identify the most relevant dimensionless groups of dynamical systems [20,21]. In other works, the Buckingham π neural network was proposed and shown to generalize better than dimensional regular neural network for predicting simple simulated physical systems, a thermal management systems in [22] and a vegetation model in [23]. Furthermore, in [22] results show that for the conducted experiments, using the Buckingham π method was more efficient than

other reduction techniques like principal component analysis and autoencoders. Finally, clustering dimensionless learning was proposed as an approach to identify the regime of behaviors [24]. All in all, many very interesting works have shown the potential of using the Buckingham π theorem to improve the generalization of machine learning models. However, since most of the results are for very simple simulated experiments and in the context of fluid mechanics prediction, many questions still remain unanswered regarding how such techniques would perform in other fields and with experimental data. In this paper, we present a case study to specifically explore the potential of such approaches in order to provide a potential solution to the problem of transfer learning for motion models of car-like vehicles.

Dimensional analysis is still very seldomly explored in the field of motion control and robotics, to our knowledge only a few projects have explored that concept. In the case of an in-lane collision avoidance system, Ref. [25] created dimensionless indices, based on a vehicle's relative longitudinal velocity, road surface conditions and terminal lateral distance, which were used in the selection of the most efficient maneuver and timing for an intervention. The use of these indices was shown to be computationally efficient. The authors of [26] proposed an improved method for kinematic calibration based on dimensionless error mapping matrices (EMMs). Simulation results show that the residual pose errors with the proposed dimensionless EMMs were lower than with the conventional EMM in various units. Finally, Ref. [27] presents the concept of dimensionless policies. It discusses the notion that similar dynamic systems must have an equivalent optimal policy when expressed in dimensionless form and shows how to exactly transfer a motion control policy. However, it does not provide results regarding the quality of transfers when the similarity condition is not exactly met, which is studied in this paper with a case study.

The contribution of this paper is to demonstrate the advantage of learning schemes leveraging the Buckingham π theorem, i.e., dimensional analysis, specifically for transferring motion models between car-like vehicles. While the idea of using dimensionless input–output variables to improve the generalization of the data-driven model is not new, the novelty of this paper is to evaluate its performance for the specific application of sharing data and models between car-like vehicles. Furthermore, it also provides an important contribution by showing that the advantage of the implemented technique was also present in a real-life experiment involving small autonomous car-like robots, as opposed to previous works where most results were based on simulated data in ideal conditions.

3. Case-Study: Predicting the Braking Behaviour of Cars

The main goal of this paper is to validate whether a learning scheme can benefit from a dimensionless mapping of its inputs and outputs. These benefits could take the form of an increase in learning speed, a reduction in mean absolute error (MAE) and the possibility of sharing knowledge between similar systems. To do so, three different learning schemes are compared using their respective MAEs; all are based on optimized distributed gradient boosting (XGBoost) models trained with simulated and experimental data. The first one, the baseline of the comparison, is using directly the XGBoost model on the physical input and output data without any prepossessing. The second one is based on transforming the input and output data into dimensionless variables, based on the Buckingham π. The third approach is also based on the Buckingham π theorem but with additional dimensionless inputs (hand-crafted features based on expert knowledge of the system's physics). Inspired by the challenges of learning good control policies for emergency braking maneuvers under various ground conditions (snow, ice, etc.), such as explored in [28], the case study used in this paper is the prediction of the final relative position of a vehicle after a sudden emergency braking maneuver, based on initial conditions, control inputs (braking and steering), and environmental parameters. The learned model could then be used in a planning and control pipeline to select the best maneuver in an emergency braking situation for instance. To simplify the case study a few assumptions are made. Firstly, the control inputs are constant throughout a maneuver, i.e., the steering angle δ and braking action

(constant deceleration rate a imposed on both rear wheels) are decided once at the start of the maneuver and remain constant until the vehicle comes to a complete stop. Also, only the final position of the vehicle is predicted, not the complete trajectory. Furthermore, the environment is always a horizontal surface and only 2D planar motion is taken into account. The simulation is based on a kinematic bicycle motion model, while for the experimental data, the dynamics of the system are taken into account, neglecting any roll or pitch effects and considering the complex soil-tire relationship as a constant coefficient of friction μ. Table 1 shows the main variables involved in this case study.

Three small-scale car-like vehicles of various dimensions are used for both validation processes (simulation and experimentation). Vehicle 1, shown in Figure 1a is a modified 1:10 scale platform based on a Traxxas SLASH. Vehicle 2, shown in Figure 1b, is also a modified 1:10 scale platform based on a Traxxas SLASH with the chassis modified to have a longer wheelbase. Finally, vehicle 3, shown in Figure 1c, is a modified 1:5 scale platform based on a Traxxas X-MAXX. They will be referred to as small vehicle, long vehicle and large vehicle respectively. Their specifications are summarized in Table 2.

Table 1. Variables used throughout this paper.

Variables	Descriptions	Units [Dimensions]
	State Variables	
X	X-axis position of the vehicle in the world frame	m [L]
Y	Y-axis position of the vehicle in the world frame	m [L]
θ	Yaw of the vehicle in the world frame	rad
	Environment related variables	
μ	Friction coefficient wheels/road	-
v	Longitudinal velocity of the vehicle	m/s [LT^{-1}]
g	Gravitational acceleration	m/s^2 [LT^{-2}]
	Maneuvers related variables	
a	Deceleration of the wheel	m/s^2 [LT^{-2}]
δ	Steering angle of front wheels	rad
	Vehicles related variables	
N_f	Normal force on front wheels	N [MLT^{-2}]
N_r	Normal force on rear wheels	N [MLT^{-2}]
l	Length between vehicle's axles	m [L]

Table 2. Specifications of the three experimental vehicles.

	Wheel Base Length	Weight on Front Wheels (Static)	Weight on Back Wheels (Static)
Vehicle	l [m]	N_f [N]	N_r [N]
Small	0.345	37.77	28.84
Long	0.853	22.74	52.89
Large	0.475	71.12	71.12

4. Learning with Simulated Data

4.1. Simulator Presentation

This section presents an analysis in which the proposed learning approach is evaluated using data generated in a simplified simulated environment. The simulation was based on the kinematic bicycle model frequently used in the literature. The non-linear equations for the bicycle model are shown in (1).

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{\theta} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} vcos(\theta) \\ vsin(\theta) \\ \frac{vtan(\delta)}{l} \\ a \end{bmatrix} \quad (1)$$

For the three vehicles presented in Figure 1, 500 simulations were performed with the initial values and control input shown in Table 3, that is, all combinations of 50 initial speed levels, 10 levels of braking, and 11 steering positions. All simulations start with the vehicle at position $[X, Y] = (0, 0)$, orientation $\theta = 0$ rad, and initial speed v_i. The maneuver starts immediately with deceleration a and steering angle δ. The simulation stops when the vehicle reaches zero velocity and the final position and orientation are noted.

Table 3. Values for the simulations for all 3 vehicles.

v_i (m/s)	a (m/s^2)	δ (rad)
From 0.1 to 5.0 by 0.1	From −0.1 g to −1.0 g by −0.1 g	From 0.0000 to 0.7854 by 0.0785
50 values	10 values	11 values

4.2. Learning Model Types for the Simulation

Here, the details of a baseline reference prediction scheme and two novel dimensionless prediction schemes are presented.

4.2.1. Traditional Dimensionalized Learning Model

The process of integrating the nonlinear equations describing the kinematic bicycle model until the vehicle comes to a stop, can be written in the form of input-relationship as given by (2).

$$\begin{bmatrix} X \\ Y \\ \theta \end{bmatrix} = f\left(v_i, a, \delta, l\right) \quad (2)$$

As a baseline learning scheme for this evaluation, the XGBoost algorithm is used directly to learn a black-box input–output model based on the data. Thus, the dimensionalized parameters for the initial velocity v_i, the deceleration a, the steering angle δ and the length of the wheelbase l are used to predict the final pose $[X, Y, \theta]$ of the vehicle. This traditional dimensionalized learning scheme is illustrated in Figure 2.

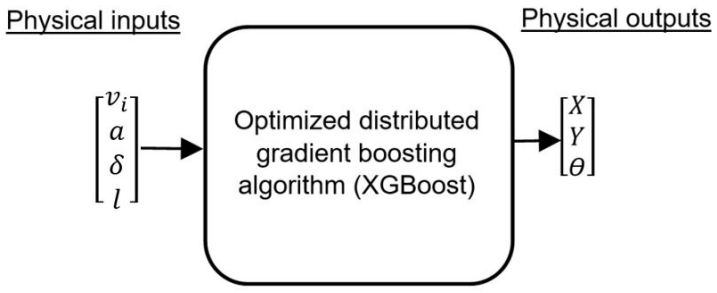

Figure 2. Baseline learning scheme for simulated data.

4.2.2. Buckingham π Theorem-Based Model

According to the Buckingham π theorem [17], a relationship with N dimensional variables and P independent dimensions, can be restated as a relationship between $m = N - P$ dimensionless numbers. Here, each line of Equation (2), represents a relationship between

N = 5 variables and P = 2 fundamental dimensions (length [L] and time [T]) involved. Each line of Equation (2) can thus be restated as a relationship between three dimensionless variables, leading to the following:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} = f\left(\pi_4, \pi_5\right) \tag{3}$$

Here, the dimensionless variables are selected using the distance between the two axles l [M] and the initial velocity of the vehicle v_i [LT^{-1}] as repeated variables in the dimensional analysis:

$$\pi_1 = \frac{X}{l} \quad \pi_2 = \frac{Y}{l} \quad \pi_3 = \theta \quad \pi_4 = \frac{al}{v_i^2} \quad \pi_5 = \delta \tag{4}$$

The new dimensionless relationship to predict is thus now restating (2) with the selected dimensionless variables, and the following input–output relationship:

$$\begin{bmatrix} \frac{X}{l} \\ \frac{Y}{l} \\ \theta \end{bmatrix} = f\left(\frac{al}{v_i^2}, \delta\right) \tag{5}$$

It is interesting to note that, in dimensionless form, the motion behavior model has a reduced input space of two dimensions instead of four initially with Equation (2). The dimensionless learning scheme consists of using the same XGBoost algorithm but on input–output data converted into those dimensionless variables. As illustrated in Figure 3, the scheme can be seen as a prepossessing of the data using intelligent scaling based on dimensional analysis.

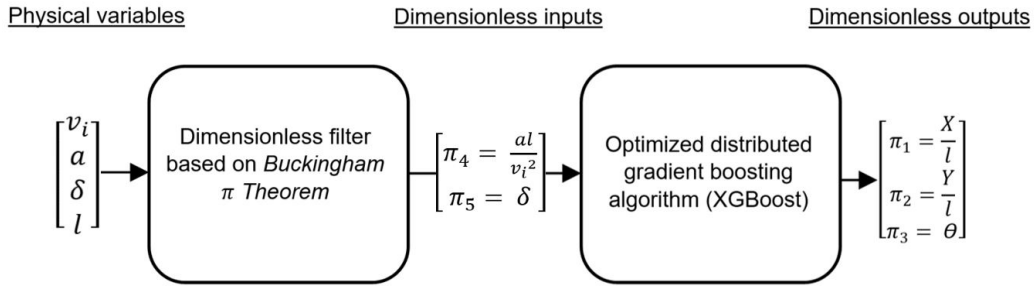

Figure 3. Buckingham π theorem-based model for simulated data.

4.2.3. Augmented Buckingham π Theorem-Based Model

A secondary question that is explored in this manuscript is the impact of adding supplementary physically relevant hand-crafted dimensionless variables as inputs to a learning algorithm. This scheme will be referred to as the Augmented Buckingham π theorem-based model. Here, an arbitrary dimensionless number π_6 that is constructed based on the initial angular velocity is included:

$$\dot{\theta} = \frac{v\tan(\delta)}{l} \qquad \Rightarrow \qquad \pi_6 = \frac{v^2\tan(\delta)}{al} \tag{6}$$

The learned input–output relationship is thus modified into the form:

$$\begin{bmatrix} \frac{X}{l} \\ \frac{Y}{l} \\ \theta \end{bmatrix} = f\left(\pi_4, \pi_5, \pi_6\right) = f\left(\frac{al}{v_i^2}, \delta, \frac{v^2\tan(\delta)}{al}\right) \tag{7}$$

The hypothesis is that adding this hand-crafted feature, which should be more closely related to predicting outputs based on first principles, might simplify the learning task and improve the performance of the learning algorithm. The inputs and outputs of the Augmented Buckingham π theorem-based model used to predict the outcome of a maneuver are shown in Figure 4.

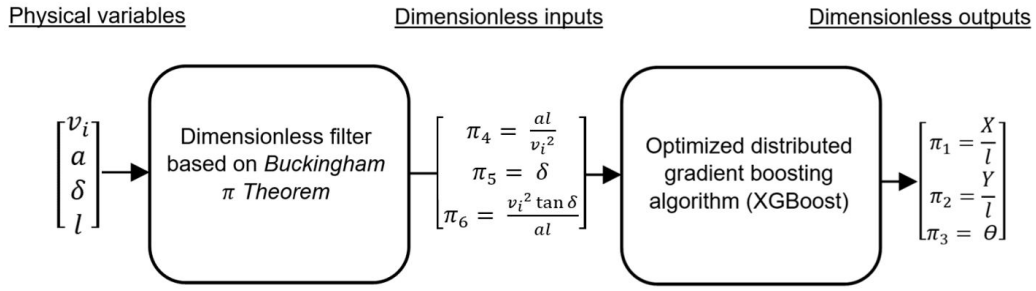

Figure 4. Augmented Buckingham π theorem-based model for simulated data.

4.3. Results

The results for each model type are examined from two perspectives: (1) the accuracy (MAEs) of self-predictions, cross-predictions, and shared predictions and (2) the effect of the size of the training set on the convergence of the MAEs, i.e., the learning rates for each situation.

4.3.1. Mean Absolute Error

The MAE is one of the most widely used metrics for regression algorithms since it has the same units as the model's dependent variables, which makes the results easier to understand. Equation (8) shows how MAE is calculated, where n is the number of tests used, y_i are the actual result values and $\hat{y}_i$ are the predicted values. Here the output values for both dimensionless models are $\pi_1 = \frac{X}{l}$, $\pi_2 = \frac{Y}{l}$ and $\pi_3 = \theta$, π_1 and π_2 are multiplied by the wheelbase l of the tested vehicle before computing the MAE so that all results share the same units.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|(y_i - \hat{y}_i)| \quad (8)$$

To evaluate the performance of the models, 80% of the simulated data was used to train a model of each type for the three vehicles, and the remaining 20% was used to test the models. Table 4 contains the results of the traditional model (baseline), and the results of the model based on the Buckingham π theorem are presented in Tables 5 and 6 show the results of the Augmented Buckingham model. For all three tables, the light gray cells correspond to the results of what we call self-predictions, i.e., predictions that use a vehicle's test data in its own trained model. The white cells correspond to the results of a vehicle's test set in another vehicle's trained model, which will be referred to as cross-predictions. Finally, the dark gray cells are the results of shared predictions. For shared predictions, the test set of a particular vehicle is run for a model that has been trained with the training data from all three vehicles (i.e., a shared database).

Table 7 summarizes the three previous Tables 4–6. For self-predictions, we can see that the Buckingham π theorem-based model and the Augmented Buckingham model are, respectively, 1.93 times and 3.60 times more precise than the traditional model. These numbers climb to 11.76 times and 15.80 times for cross-predictions. Finally, the Buckingham π theorem-based model and the Augmented Buckingham model are, respectively, 4.80 times and 9.17 times more precise than the traditional model in shared-predictions.

Table 4. MAEs for the traditional baseline model (X and Y are in meters and θ in rad). Light gray cells correspond to self-predictions, white cells correspond to cross-predictions, and dark gray cells are predictions from merged data.

Models	Data Vehicle 1 (Small)	Data Vehicle 2 (Long)	Data Vehicle 3 (Large)
Model Vehicle 1 (Small)	X: 0.0494 Y: 0.0466 θ: 0.0587	X: 0.3703 Y: 0.3071 θ: 0.9851	X: 0.1601 Y: 0.1459 θ: 0.4502
Model Vehicle 2 (Long)	X: 0.3653 Y: 0.3032 θ: 0.9919	X: 0.0316 Y: 0.0257 θ: 0.0238	X: 0.2622 Y: 0.2061 θ: 0.5369
Model Vehicle 3 (Large)	X: 0.1667 Y: 0.1440 θ: 0.4599	X: 0.2679 Y: 0.2125 θ: 0.5324	X: 0.0414 Y: 0.0362 θ: 0.0422
Model Merged (All 3)	X: 0.0467 Y: 0.0465 θ: 0.0457	X: 0.0433 Y: 0.0368 θ: 0.0302	X: 0.0514 Y: 0.0498 θ: 0.0431

Table 5. MAEs for the Buckingham π theorem-based model (X and Y are in meters and θ in rad). Light gray cells correspond to self-predictions, white cells correspond to cross-predictions, and dark gray cells are predictions from merged data.

Models	Data Vehicle 1 (Small)	Data Vehicle 2 (Long)	Data Vehicle 3 (Large)
Model Vehicle 1 (Small)	X: 0.0204 Y: 0.0261 θ: 0.0331	X: 0.0144 Y: 0.0102 θ: 0.0109	X: 0.0125 Y: 0.0099 θ: 0.0159
Model Vehicle 2 (Long)	X: 0.0354 Y: 0.0394 θ: 0.2007	X: 0.0145 Y: 0.0176 θ: 0.0127	X: 0.0303 Y: 0.0262 θ: 0.0917
Model Vehicle 3 (Large)	X: 0.0153 Y: 0.0162 θ: 0.0666	X: 0.0103 Y: 0.0095 θ: 0.0097	X: 0.0190 Y: 0.0182 θ: 0.0241
Model Merged (All 3)	X: 0.0080 Y: 0.0086 θ: 0.0143	X: 0.0087 Y: 0.0087 θ: 0.0080	X: 0.0083 Y: 0.0083 θ: 0.0115

Table 6. MAEs for the Augmented Buckingham π theorem-based model (X and Y are in meters and θ in rad). Light gray cells correspond to self-predictions, white cells correspond to cross-predictions, and dark gray cells are predictions from merged data.

Models	Data Vehicle 1 (Small)	Data Vehicle 2 (Long)	Data Vehicle 3 (Large)
Model Vehicle 1 (Small)	X: 0.0126 Y: 0.0121 θ: 0.0131	X: 0.0117 Y: 0.0090 θ: 0.0043	X: 0.0104 Y: 0.0077 θ: 0.0063
Model Vehicle 2 (Long)	X: 0.0271 Y: 0.0286 θ: 0.1407	X: 0.0103 Y: 0.0134 θ: 0.0052	X: 0.0219 Y: 0.0222 θ: 0.0591
Model Vehicle 3 (Large)	X: 0.0130 Y: 0.0133 θ: 0.0389	X: 0.0090 Y: 0.0091 θ: 0.0038	X: 0.0117 Y: 0.0129 θ: 0.0098
Model Merged (All 3)	X: 0.0045 Y: 0.0050 θ: 0.0060	X: 0.0055 Y: 0.0061 θ: 0.0028	X: 0.0046 Y: 0.0051 θ: 0.0035

Table 7. MAEs comparison for all models trained with simulated data (X and Y are in meters and θ in rad). Light gray cells correspond to self-predictions, white cells correspond to cross-predictions, and dark gray cells are predictions from merged data.

Prediction Type	Traditional Dimensionalized Model	Buckingham π Theorem Based Model	Augmented Buckingham π Model
Self Predictions	X: 0.0408	X: 0.0180	X: 0.0115
	Y: 0.0362	Y: 0.0206	Y: 0.0128
	θ: 0.0416	θ: 0.0233	θ: 0.0094
Cross Predictions	X: 0.2654	X: 0.0197	X: 0.0155
	Y: 0.2198	Y: 0.0186	Y: 0.0128
	θ: 0.6594	θ: 0.0659	θ: 0.0422
Shared Predictions	X: 0.0471	X: 0.0083	X: 0.0049
	Y: 0.0444	Y: 0.0085	Y: 0.0054
	θ: 0.0397	θ: 0.0113	θ: 0.0041

4.3.2. Learning Rate

The amount of data required to obtain an operational predictive function for self-predictions for all model types is also analyzed. Figure 5 shows the MAE convergence for the final longitudinal position X, final lateral position Y and final yaw θ of the three model types with the large vehicle's self-predictions. For all outputs, the Augmented Buckingham model converges the fastest, followed by the Buckingham π theorem-based model. More data would be required to achieve full convergence of the final X and Y positions for the traditional dimensionalized model. Only the data for the large vehicle are presented for shortness, as the learning curve for all vehicles was similar.

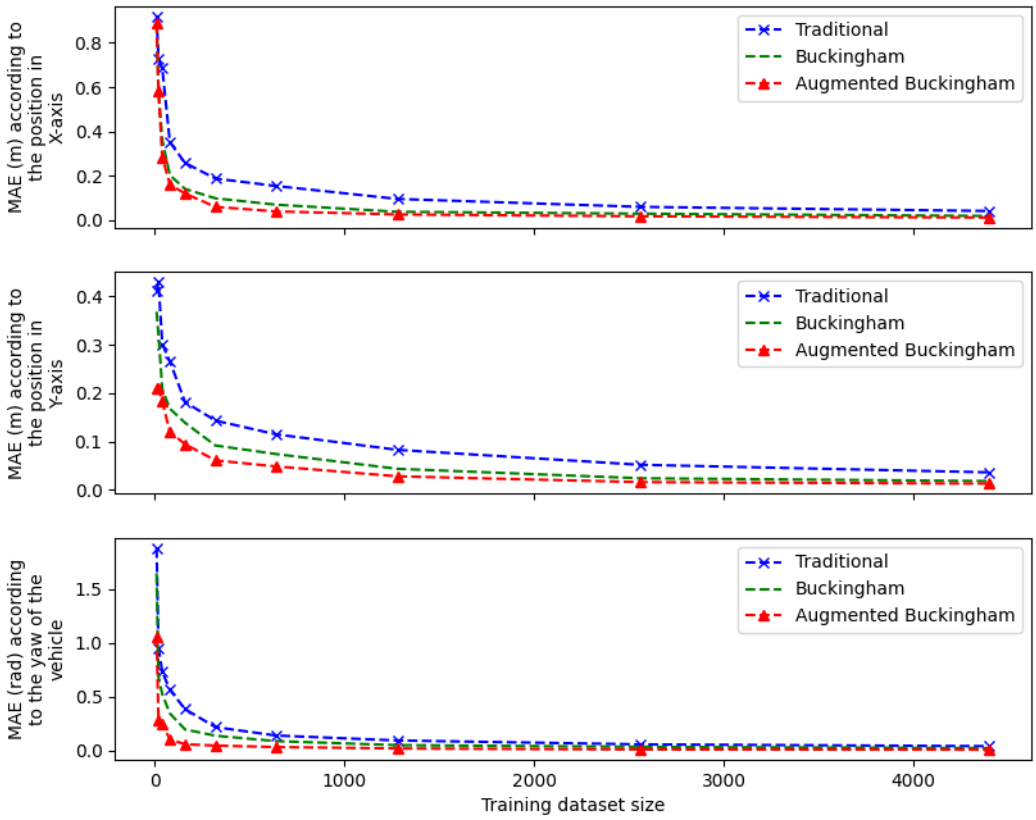

Figure 5. MAE convergence according to training size for the simulated data of the large vehicle.

4.3.3. Comparative Analysis

Various hypotheses could explain the performance improvement. (1) Normalization: It is possible that the improvement is mainly due to a good normalization of the range of the input–output variables which helps the learning algorithm. (2) Dimensionality reduction: It is possible that the improvement is mainly due to the reduction in the number of inputs. (3) Better coordinate system: It is possible that the improvement is mainly due to the coordinate transformation in the dimensionless space that makes learning easier and the knowledge more generic. In order to better understand the results and the source of improvements, a comparative analysis is conducted with multiple variations to help isolate the main source of improvements. The results are summarized in Figure 6.

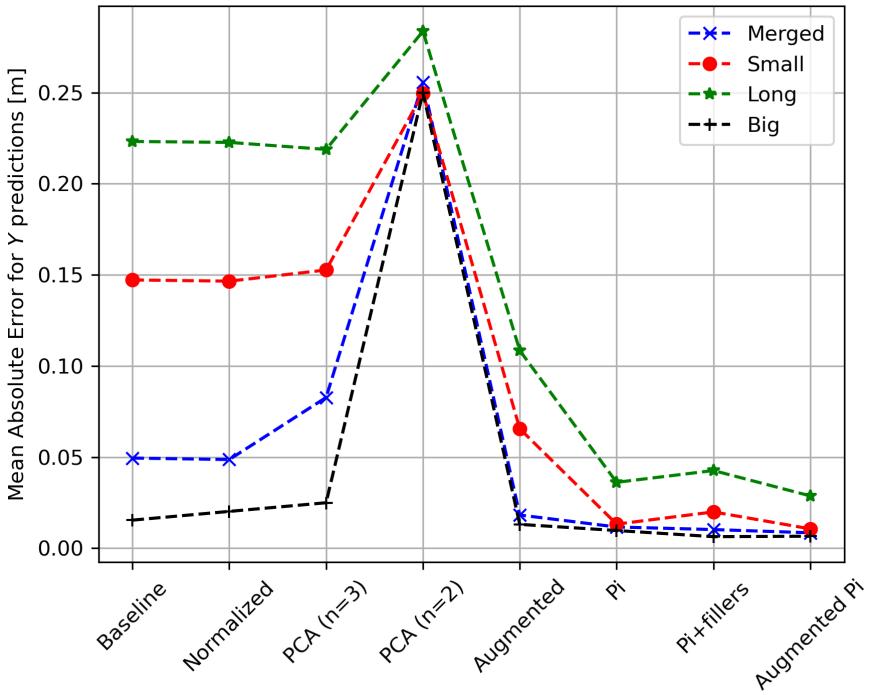

Figure 6. Comparison of the performance of various pre-processing schemes for predicting the Y position of the large vehicle. The x-axis indicates the pre-processing scheme used on the data before training the XGBoost algorithm. The four curves represent the source of the data used for the training, as indicated in the legend. The green and red curves correspond to transfer learning tasks, i.e., predicting the motion of the large vehicle using only data from other vehicles.

For this analysis, the learning problem is simplified to predicting only one output: the final Y position of the large vehicle. The performance of each learning scheme will be compared in terms of the MAEs of the prediction for a test dataset. In addition to the three schemes already presented, i.e., the baseline, the Buckingham π and the augmented Buckingham π, five new variations are compared. Normalized: The data are first normalized by the largest value seen in the training data (for each input variable). PCA: A principal component analysis (PCA) transformation is conducted on the input data before training the model and making predictions. The results include when three or two components are used, with two corresponding to the number of inputs when using the Buckingham π scheme. Augmented: The input data are augmented with a handcrafted feature, corresponding to $\dot{\theta}$ at Equation (6), i.e., the same as the Augmented Buckingham π scheme but without the additional step to make it a dimensionless number. Pi + fillers: The two π groups of the Buckingham π scheme are augmented with two original dimensional input variables: v_i and l, i.e., superfluous information according to dimensional analysis.

The results illustrated in Figure 6, show that compared to the baseline approach, normalizing the data has almost no effect on the performance, PCA reduction in the in-

put data always leads to a worse performance, and augmenting the data with $\dot{\theta}$ always leads to a good performance improvement, but less than when using the Buckingham π-based schemes. Finally, comparing the three Buckingham-based schemes, augmenting the data can increase or decrease the performance, according to the relevance of the additional features.

4.3.4. Discussion

As was expected, the dimensionless models performed much better for generalizing the predictions, i.e., for predicting the motion of a vehicle using data from a single different vehicle (cross-predictions), or a shared database of the three vehicles (shared predictions). It is also interesting, even surprising, that there was an improvement in the accuracy of the self-prediction, for which the benefit of using dimensionless variables is less direct. From the comparative study, it seems that there is something fundamentally better with using dimensionless variables as the coordinate system for encoding the input information since alternate preprocessing schemes did not lead to similar improvements.

Regarding the Augmented Buckingham model, there was a significant improvement compared to the Buckingham model without the hand-crafted input variable. The improvement can be seen mainly early in the learning phase (see Figure 5), but then as more data are coming, the performance of the other models is slowly catching up. A hypothesis is that this model has a head-start, already knowing some of the non-linearity (the tan term) while the other has to learn it from the data. All in all, this experiment is interesting because all the variables involved in generating the data were exactly known. This illustrated how drastic the improvement could be in a perfect situation by leveraging dimensional analysis in the learning scheme.

5. Learning with Experimental Data

The validation presented in the last section consisted of a very clean ideal situation. Here, another validation is conducted for a much more challenging learning task. Here, the task is to predict much more dynamic braking maneuvers of real small platforms, based on real data recorded with a vision-based acquisition system to provide the position of the vehicles. Hence, the problem is harder because (1) the data are noisy, (2) the behavior to model is more complex and (3) all the relevant variables are not known, and some simplifying hypothesis must be made.

5.1. Methodology for Generating the Data

For all three vehicles, 540 experimental tests were carried out with the initial values and maneuvers shown in Table 8: 3 types of terrains, 6 initial velocity levels, 10 levels of braking and 3 steering angles. All the tests are conducted indoors on a flat and uniform surface, with various material to modify the friction coefficient. All tests start with the vehicle in position $(0, 0)$, orientation 0 rad, initial speed v_i and ground/tire friction coefficient μ. The maneuver begins immediately with a deceleration of the rear wheel given by a (the servo-motor is used to impose an angular deceleration of a/r where r is the wheel radius) and setting the steering angle to position δ. The test stops when the vehicle has come to a complete stop, and the final position and orientation are recorded using a VICON triangulation camera system [29].

Table 8. Input values for the experimental t for all three vehicles.

μ	v_i (m/s)	a g (9.81 m/s^2)	δ (Rad)
0.2	From 1.0	From 0.1	0.0000
0.4	to 3.5	to 1.0	0.3927
0.9	by 0.5	by 0.1	0.7854
3 values	6 values	10 values	3 values

5.2. Model for the Dynamic Maneuvers

To capture some of the dynamic behavior of the dynamic maneuvers, more input variables are considered in the learned predictor function here. The relevant variable to include were sectioned with the assumption that a dynamic bicycle model (often used in planing and control pipeline of autonomous robots [2]) would be a good approximation of the behaviour. Additionally, the tire–ground interaction is assumed to be captured by the constant friction coefficient μ throught the maneuver. Hence, all the states, inputs and vehicle parameters that appear in those equations are used to form the input of the predictor model in this section, see Table 1.

5.2.1. Traditional Baseline Model

The process of integrating the nonlinear equations describing the dynamic bicycle model until the vehicle comes to a stop can be written in the following input-relationship form:

$$\begin{bmatrix} X \\ Y \\ \theta \end{bmatrix} = f\left(\mu, v_i, g, a, \delta, N_f, N_r, l\right) \tag{9}$$

The learned model for the dynamic maneuver will thus be based on using the same input variables to predict the final position obtained for the real experimental maneuver. Here, it is important to note that this will be an approximation: some variables that might have some influence on the outcome (for instance the suspension rigidity) are not included as available inputs for the model. The baseline learning scheme for the experimental validation is, as illustrated in Figure 7, using directly the XGBoost algorithm to learn a data-driven approximation to Equation (9) relationship.

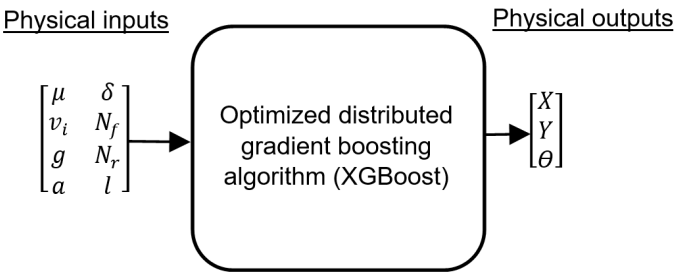

Figure 7. Baseline learning scheme for experimental data.

5.2.2. Buckingham π Theorem-Based Model

Here, if we conduct the dimensional analysis, each line of the relationship of Equation (9) involves $N = 9$ variables and $P = 3$ independent dimensions (mass [M], length [L] and times [T]). Then, based on the Buckingham π theorem, each relationship (lines of Equation (9)) could be restated as the relationship between six dimensionless groups, as follows:

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} = f(\pi_4, \pi_5, \pi_6, \pi_7, \pi_8) \tag{10}$$

By using the distance between the two axles l [M], the initial velocity of the vehicle v_i [MT^{-1}] and the normal force on front wheels N_f [MLT^{-2}] as repeated variables, the following eight dimensionless numbers were used:

$$\pi_1 = \frac{X}{l} \quad \pi_2 = \frac{Y}{l} \quad \pi_3 = \theta \quad \pi_4 = \frac{al}{v_i^2} \tag{11}$$

$$\pi_5 = \delta \quad \pi_6 = \frac{N_f}{N_r} \quad \pi_7 = \mu \quad \pi_8 = \frac{gl}{v_i^2} \tag{12}$$

The new dimensionless relationship to predict is thus in the following form:

$$\begin{bmatrix} \frac{X}{l} \\ \frac{Y}{l} \\ \theta \end{bmatrix} = f\left(\frac{al}{v_i^2}, \delta, \frac{N_f}{N_r}, \mu, \frac{gl}{v_i^2}\right) \tag{13}$$

and the overall pipeline for making predictions with the dimensionless approach is shown in Figure 8.

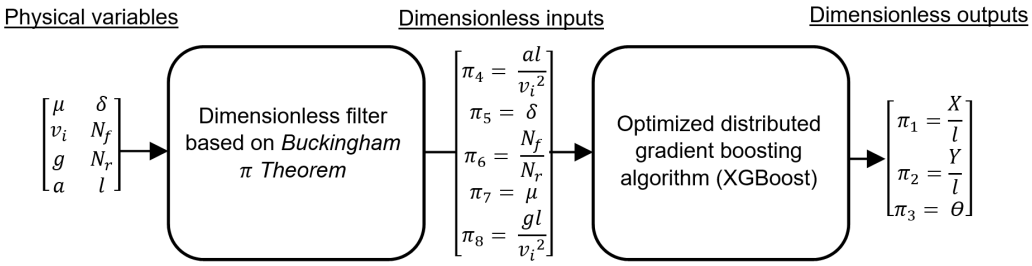

Figure 8. Buckingham π theorem-based model for experimental data.

5.2.3. Augmented Buckingham π Theorem-Based Model for Experimental Data

Here, to test if a handcrafted feature would be advantageous, two additional dimensionless numbers are chosen for their dynamic significance. Two ratios representing adherence limits for the longitudinal force π_9 and the lateral force π_{10} are introduced as described by Equations (14) and (15):

$$\pi_9 = \frac{F_{limit}}{F_{x_{maneuver}}} = \frac{N_r \mu}{ma} = \frac{N_r \mu g}{(N_f + N_r)a} \tag{14}$$

$$\pi_{10} = \frac{F_{limit}}{F_{y_{maneuver}}} = \frac{N_r \mu}{\frac{mv^2}{R}} = \frac{g\mu R}{v^2} = \frac{g\mu l}{v^2 \tan \delta} \tag{15}$$

The new dimensionless relationship to predict is thus in the following form:

$$\begin{bmatrix} \frac{X}{l} \\ \frac{Y}{l} \\ \theta \end{bmatrix} = f\left(\frac{al}{v_i^2}, \delta, \frac{N_f}{N_r}, \mu, \frac{gl}{v_i^2}, \frac{N_r \mu g}{(N_f + N_r)a}, \frac{g\mu l}{v^2 \tan \delta}\right) \tag{16}$$

and the overall pipeline for making predictions with the dimensionless approach is shown in Figure 9.

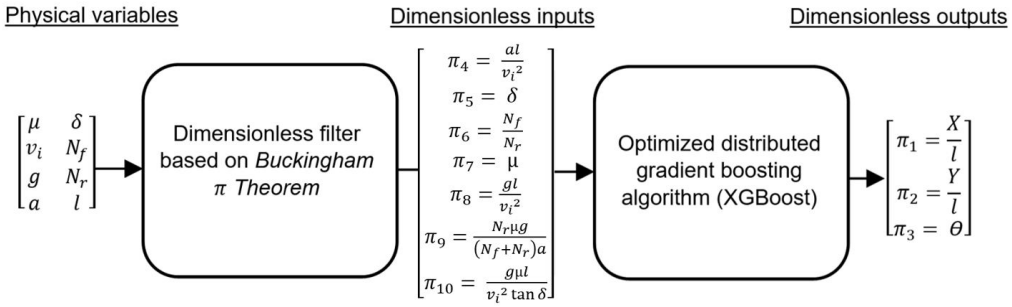

Figure 9. Augmented Buckingham π theorem-based model for experimental data.

5.3. Results

As for the simulation data, the results for each model type are examined from two perspectives: (1) the accuracy (MAEs) and (2) the learning rates for each situation.

5.3.1. Mean Absolute Error

To avoid any redundancy, only the summarizing Table 9 for self, cross and shared predictions is presented for the experimental data. For self-predictions, we can see that the Buckingham π theorem-based model and the Augmented Buckingham model are, respectively, 1.11 times and 1.14 times more precise than the traditional model on average. These numbers drop to 1.05 times and 1.07 times more precise for cross-predictions. Finally, the Buckingham π theorem-based model and the Augmented Buckingham model are, respectively, 1.11 times and 1.28 times more precise than the traditional model in shared-predictions.

Table 9. MAE comparison for all models trained with experimental data (X and Y are in meters and θ in rad). Light gray cells correspond to self-predictions, white cells correspond to cross-predictions, and dark gray cells are predictions from merged data.

Prediction Types	Traditional Dimensionalized Model	Buckingham π Theorem Based Model	Augmented Buckingham π Model
Self Predictions	X: 0.1229	X: 0.1228	X: 0.1197
	Y: 0.0488	Y: 0.0321	Y: 0.0335
	θ: 0.0482	θ: 0.0596	θ: 0.0554
Cross Predictions	X: 0.3027	X: 0.3752	X: 0.3775
	Y: 0.2198	Y: 0.0186	Y: 0.0128
	θ: 0.2007	θ: 0.1494	θ: 0.1454
Shared Predictions	X: 0.0539	X: 0.0507	X: 0.0400
	Y: 0.0130	Y: 0.0105	Y: 0.0089
	θ: 0.0193	θ: 0.0187	θ: 0.0187

5.3.2. Learning Rate

Figure 10 shows the MAE convergence for the final longitudinal position X, final lateral position Y and final yaw θ of the three model types with the large vehicle's experimental self-predictions. For all outputs, all three models follow a similar learning curve. More training data (limited here by the number of experimental tests conducted) would be required to fully converge since all curves keep on improving at a relatively high rate.

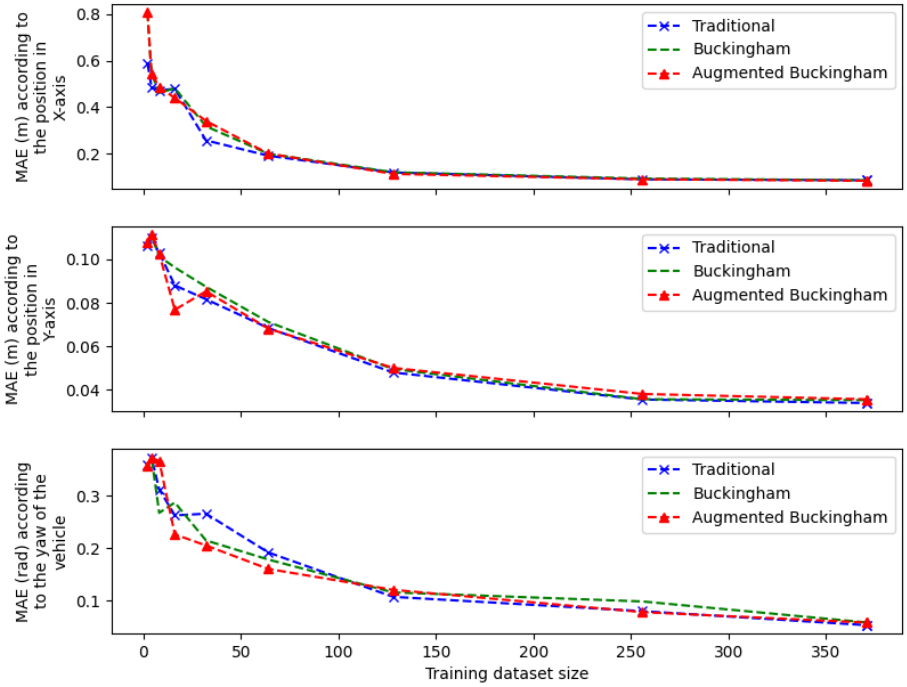

Figure 10. MAE convergence according to training size for the experimental data of the large vehicle.

5.3.3. Discussion

Here, it is interesting to note that even for the much harder realistic learning task, the proposed Buckingham π schemes still led to always improving the prediction accuracy when using a shared database from the data of the three vehicles. However, the improvements were found to be less spectacular than for the clean simulated data. This was expected since for this experiment: (1) the data were noisy (the main source would be the vision-based measurement system), (2) the input–output relationship to predict is more complex (mainly the complex tire-ground relationship), and (3) the given list of relevant inputs to the model is not a complete list of variables that influence the outcome (unlike previously when the data were generated by a known controlled simulation model), it was only an approximation based on the dynamic bicycle model equations. In future works, it would be interesting to conduct more intermediary experiments to isolate the effect of each difficulty individually.

6. Conclusions

To conclude, this paper presents a dimensionless learning scheme designed to allow the sharing of data and models between cars of various kinematic and dynamic parameters and results from a case study demonstrating the performance of the scheme to generalize in comparison to a baseline learning approach. The results of the case study based on kinematic simulations show that the proposed dimensionless models are much more accurate when using a shared database (an order of magnitude more accurate), and have faster learning rates than a traditional dimensional model. The results of the case study based on experimental data of car-like robots conducting dynamic braking maneuvers also showed an advantage with the proposed dimensionless scheme over a baseline approach (in the range of 5% to 30% of precision improvements). The gain in the experimental context was less significant than what was observed through simulations, which is mainly attributable to a discrepancy between the assumptions made (i.e., when selecting the list of input variables for the dimensional analysis) and the high complexity of the tire dynamics involved in the experiments. However, it is interesting to note that using a dimensionless relationship for the predictions always led to improvements in all the tested scenarios. Since it is an easy prepossessing step to conduct it should be used all the time, especially since if all relevant variables are included this is a lossless technique to reduce the dimension of the input and output space according to the Buckingham π theorem. In future work, it would be interesting to conduct more tests in intermediary situations to better map the potential improvements one could expect by using the presented dimensionless approach.

All in all, the results presented in this case study have shown that leveraging Buckingham π theorem can lead to very interesting improvements in terms of experimental data required to train a learning algorithm modeling a vehicle, to transfer results between similar vehicles, and when using a shared database. Since the requirements for using dimensional analysis and applying the Buckingham π theorem are simply that the relationship (to predict) involves physically meaningful variables and knowing which one, we should expect similar results for generalizing a learned model involving various types of vehicles, robotics systems, or in fact any physically meaningful non-linear relationship to predict. However, there are still many unanswered questions regarding this type of learning technique. Hence, the approach should be investigated further to better understand its full potential and limitations.

Author Contributions: Conceptualization, W.T., O.L. and A.G.; methodology, W.T., O.L. and A.G.; software, W.T.; validation, W.T. and O.L.; formal analysis, W.T. and A.G.; investigation, W.T., O.L. and A.G.; resources, W.T., O.L. and A.G.; data curation, W.T.; writing—original draft preparation, W.T., O.L. and A.G.; writing—review and editing, A.G.; visualization, W.T.; supervision, A.G.; project administration: A.G.; funding acquisition, A.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSERC discovery grant number RGPIN-2018-05388.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MAS	Mean Absolute Error
PCA	Principal Component Analysis

References

1. Amer, N.H.; Zamzuri, H.; Hudha, K.; Kadir, Z.A. Modelling and Control Strategies in Path Tracking Control for Autonomous Ground Vehicles: A Review of State of the Art and Challenges. *J. Intell. Robot. Syst.* **2017**, *86*, 225–254. [CrossRef]
2. Paden, B.; Čáp, M.; Yong, S.Z.; Yershov, D.; Frazzoli, E. A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles. *IEEE Trans. Intell. Veh.* **2016**, *1*, 33–55. [CrossRef]
3. Katrakazas, C.; Quddus, M.; Chen, W.H.; Deka, L. Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions. *Transp. Res. Part C Emerg. Technol.* **2015**, *60*, 416–442. [CrossRef]
4. Aradi, S. Survey of Deep Reinforcement Learning for Motion Planning of Autonomous Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2020**, *23*, 740–759. [CrossRef]
5. Crites, R.H.; Barto, A.G. Improving Elevator Performance Using Reinforcement Learning. *Adv. Neural Inf. Process. Syst.* **1995**, *8*, 1017–1023.
6. Di, X.; Shi, R. A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to AI-guided driving policy learning. *Transp. Res. Part C Emerg. Technol.* **2021**, *125*, 103008. [CrossRef]
7. Lake, B.M.; Ullman, T.D.; Tenenbaum, J.B.; Gershman, S.J. *Building Machines That Learn and Think Like People*; Cambridge University Press: Cambridge, UK, 2017; Volume 40. [CrossRef]
8. Liu, Y.; Xu, H.; Liu, D.; Wang, L. A digital twin-based sim-to-real transfer for deep reinforcement learning-enabled industrial robot grasping. *Robot. Comput.-Integr. Manuf.* **2022**, *78*, 102365. [CrossRef]
9. Dey, S.; Boughorbel, S.; Schilling, A.F. Learning a Shared Model for Motorized Prosthetic Joints to Predict Ankle-Joint Motion. *arXiv* **2021**, arXiv:2111.07419.
10. Andrychowicz, O.M.; Baker, B.; Chociej, M.; Józefowicz, R.; McGrew, B.; Pachocki, J.; Petron, A.; Plappert, M.; Powell, G.; Ray, A.; et al. Learning dexterous in-hand manipulation. *Int. J. Robot. Res.* **2020**, *39*, 3–20. [CrossRef]
11. Nagabandi, A.; Clavera, I.; Liu, S.; Fearing, R.S.; Abbeel, P.; Levine, S.; Finn, C. Learning to Adapt in Dynamic, Real-World Environments Through Meta-Reinforcement Learning. *arXiv* **2019**, arXiv:1803.11347.
12. Dasari, S.; Ebert, F.; Tian, S.; Nair, S.; Bucher, B.; Schmeckpeper, K.; Singh, S.; Levine, S.; Finn, C. RoboNet: Large-Scale Multi-Robot Learning. *arXiv* **2020**, arXiv:1910.11215.
13. Sorocky, M.J.; Zhou, S.; Schoellig, A.P. Experience Selection Using Dynamics Similarity for Efficient Multi-Source Transfer Learning Between Robots. *arXiv* **2020**, arXiv:2003.13150.
14. Chen, T.; Murali, A.; Gupta, A. Hardware Conditioned Policies for Multi-Robot Transfer Learning. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 9355–9366.
15. Bertrand, J. Sur l'homogénéité dans les formules de physique. *Cah. Rech. l'Acad. Sci.* **1878**, *86*, 916–920.
16. Rayleigh, L. VIII. On the question of the stability of the flow of fluids. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1892**, *34*, 59–70. [CrossRef]
17. Buckingham, M.E. On Physically Similar Systems; Illustrations of the Use of Dimensional Equations. *Phys. Rev.* **1914**, *4*, 345–376. [CrossRef]
18. Fukami, K.; Taira, K. Robust machine learning of turbulence through generalized Buckingham Pi-inspired pre-processing of training data. In Proceedings of the APS Division of Fluid Dynamics Meeting Abstracts, Phoenix, AZ, USA, 21–23 November 2021; p. A31.004.
19. Fukami, K.; Goto, S.; Taira, K. Data-driven nonlinear turbulent flow scaling with Buckingham Pi variables. *J. Fluid Mech.* **2024**, *984*, R4. [CrossRef]
20. Bakarji, J.; Callaham, J.; Brunton, S.L.; Kutz, J.N. Dimensionally consistent learning with Buckingham Pi. *Nat. Comput. Sci.* **2022**, *2*, 834–844. [CrossRef]
21. Xie, X.; Samaei, A.; Guo, J.; Liu, W.K.; Gan, Z. Data-driven discovery of dimensionless numbers and governing laws from scarce measurements. *Nat. Commun.* **2022**, *13*, 7562. [CrossRef] [PubMed]
22. Oppenheimer, M.W.; Doman, D.B.; Merrick, J.D. Multi-scale physics-informed machine learning using the Buckingham Pi theorem. *J. Comput. Phys.* **2023**, *474*, 111810. [CrossRef]
23. Villar, S.; Yao, W.; Hogg, D.W.; Blum-Smith, B.; Dumitrascu, B. Dimensionless machine learning: Imposing exact units equivariance. *arXiv* **2022**, arXiv:2204.00887.

24. Zhang, L.; Xu, Z.; Wang, S.; He, G. Clustering dimensionless learning for multiple-physical-regime systems. *Comput. Methods Appl. Mech. Eng.* **2024**, *420*, 116728. [CrossRef]
25. Singh, A.S.P.; Osamu, N. Nondimensionalized indices for collision avoidance based on optimal control theory. In Proceedings of the 36th FISITA World Automotive Congress, Busan, Republic of Korea, 26–30 September 2016.
26. Luo, X.; Xie, F.; Liu, X.J.; Xie, Z. Kinematic calibration of a 5-axis parallel machining robot based on dimensionless error mapping matrix. *Robot. Comput.-Integr. Manuf.* **2021**, *70*, 102115. [CrossRef]
27. Girard, A. Dimensionless Policies Based on the Buckingham Pi Theorem: Is This a Good Way to Generalize Numerical Results? *Mathematics* **2024**, *12*, 709. [CrossRef]
28. Lecompte, O.; Therrien, W.; Girard, A. Experimental Investigation of a Maneuver Selection Algorithm for Vehicles in Low Adhesion Conditions. *Trans. Intell. Veh.* **2022**, *7*, 407–412. [CrossRef]
29. Vicon. Available online: https://www.vicon.com/ (accessed on 20 May 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.