Technology Convergence 2023
Setting the Wheels In Motion: Reimagining the future of heavy vehicles, roads and freight

Designing high speed weigh-in-motion system with local linear regression in Wallonia (Belgium)
Towards direct weight enforcement

Loïc Warscotte
Researcher,
University of
Namur, Belgium.

Jehan Boreux, PhD
Walloon Public Service,
Belgium

Abstract

The use of High Speed Weigh-In-Motion (HS-WIM) towards direct weight enforcement is growing worldwide. Weighing sensors can be implemented inside the roadway and directly estimated the weight of the moving vehicle. Regarding the specific system in Wallonia (Belgium), the class A (according to the COST323) and the class I (according to the E1318-94) are reached for the trucks T2S3.
However, this accuracy is only obtained for a small number of vehicles, which are filtered by using an algorithm of selection to filter which vehicles to weigh. The purpose of this paper is to present an algorithm with the capability of increasing the number of selected vehicles without decreasing the good accuracy of the raw system. This work aims to compare different mathematical estimators and choose the best hyper-parameters by cross-validation. We especially bring up the topic of the accuracy of the HS-WIM for each estimator regarding the number of needed vehicles to train our model. We conclude that the local regression seems the more suitable estimator in many aspects, that we describe in the following. Finally, the ability of the model to be implemented (which are mainly the time and space complexity) in a electronic device is also discussed.

Keywords: HS-WIM, Weighing, Metrology, Accuracy, Local linear regression

1. Introduction

Many issues as a deterioration of the roadway or a lack of safety may arise due to overloaded vehicles on a highway. Controlling and punishing the drivers is a heavy work because each vehicle must be weighed off the highway, with the nearest static weighing machine. This approach requires a lot of time and only a small number of the vehicles can be weighed.
Nowadays, weighing sensors are able to directly weigh vehicles on the road in order to save time and control a much bigger number of vehicles. Then, the owner of the vehicle is given a penalty if the estimated weight is beyond the allowed range.

In the last few years, High Speed Weigh-In-Motion (HS-WIM) gained attention in civil engineering. Indeed, papers as (Jacob, 2016), (Gadja, 2016), (Burnos, 2017) and (Gadja, 2023) described the behavior of the sensors regarding the external conditions as temperature of the road or the speed of the vehicles.
Moreover, technical papers as (Jacob, 2000) established the practical protocols for the HS-WIM towards direct enforcement. These papers give information about the statistical background and the procedure to verify the accuracy of a HS-WIM system. They also provide a classification of the sites according to the quality and the specifications of the pavement. Because of the general growing enthusiasm for automatic weight enforcement, the COST323 (COST, 1998) and the E1318 (ASTM, 2021) which are devoted to the general recommendations of a Weigh-In-Motion (WIM) system, tends to become European and U.S. pre-standard. Some papers as (Antofie, 2019), (Doupal, 2016) and (Marchadour, 2008) are dedicated to the description of the accuracy that a specific system is able to reach by using the HS-WIM towards direct enforcement. Besides, papers as (Cebon, 1991), (Stergioulas, 2000) and recently (Gadja, 2020) estimate the static forces from the total forces applied by the vehicles on the road with the used of multiple sensors.

In Wallonia (Belgium), the Walloon Public Service (SPW) and the compagny Sterela start a partnership to install and calibrate quartz sensors at several places in 2019 in order to face the problem of overloading vehicles. These sensors were installed at the same time as a re-foundation of the roadway. They are able to record a lot of information about the vehicles, such as the lateral position of the wheels on the road, the speed or the weight. The figure 1 shows the magnetics loops, the weighing sensors, the temperature sensors and the sensors of position represented by the dotted rectangle, the hatched rectangles, the dots and the lines, respectively. The reached accuracy of this specific raw system is very high, thanks to the statistical and physical models designed by Sterela. According to the COST323 (and E1318) specifications, the system currently belongs to the class A (and class I) for the trucks T2S3 and to the class B (and class I) for the vans U2.

Whereas the hardware of Sterela provides raw estimations thanks to the piezoelectricity equations, the dynamical effect of the high speed on the axles typically leads to high estimation errors for a lot of vehicles. As a consequence, an important selection of vehicles is needed to obtain a good accuracy in practice, which are detailed in (Antofie, 2019).
This selection consists of multiple conditions which must be met in order to give an estimation of the weight with high accuracy. If this selection is passed for a vehicle, its validity state is said to be 1 and the estimated weight is accepted, and 0 otherwise.

Instead of discarding the unusual vehicles which failed the selection, the purpose of this paper is to introduce a suitable model which aims to predict the weight with a high precision and acceptance rate without having to pass a hard selection.

While most of papers are concerned with the description of the accuracy, much less researches are dedicated to the development of a mathematical model to improve it, or a processing of the high dimensional data. For instance, (Burnos, 2021) and (Burnos, 2020) proposed in Poland an algorithm to update parameters in the computation of weights according to the time of the day or the season. This updating method is valuable if some variables rapidly change and is completely independent of the static weight, as the temperature in Poland which can vary over a range of dozens of degrees in few hours.

Concerning high parametric mathematical models, (Heidari, 2019) proposed auto-encoders and regressors with feed-forward networks. As expected, these networks outperform the raw system. Nevertheless, they need thousands of vehicles to train the networks. In Belgium, much less time is dedicated to the creation of a database and this large number of weighing on a low speed scale is impossible. At first sight, training a high-dimensional model with a large number of parameters seems appealing to highlight non-linear relationships between the features of a single vehicle. Though, these high-parametric methods easily suffer from over-fitting and high sensitivity which result in bad accuracy for small or noisy training sets (Bishop, 2006).

More sparse models as extended linear regression have been implemented in the case of on-board WIM in (Kirushanth, 2020). Indeed, combinations and non-linear transformations of the existing variables were performed, such as power and multiplications of variables, or application of log functions. Afterwards, the more important features was selected thanks to a stepwise feature selection. This procedure is well suited in the case of a data set of moderate size. However, assuming some hypothesis about the non-linear functions are sometimes difficult and leads to less accurate estimations. Moreover, the validation step might highly depend on the site and had then to be performed at each location.

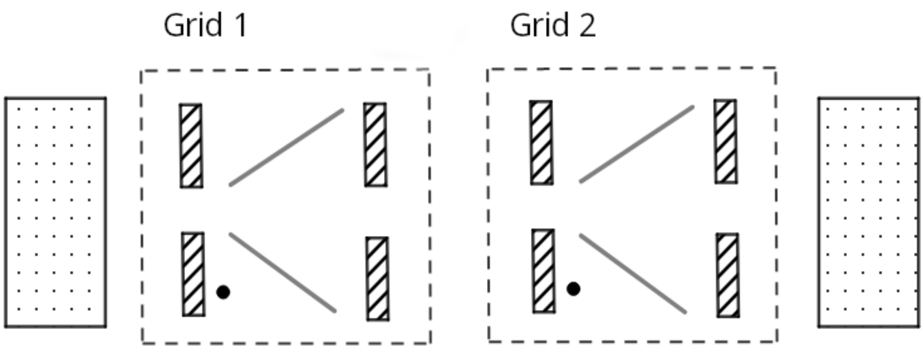

Figure 1 – Scheme of the two grids of sensors in the roadway

In order to face our problem, the following procedure designed by the SPW in this partnership aims to model the residual errors made which are not corrected by the sensors. In this case, the challenging task is having a balance between the accuracy of the weighing and the needed size of the training set, due to the difficulty of the data collecting.
Therefore, the first step of this procedure is to handle the high-dimensionality of the data with a reduction of dimension and then developing a model that takes the reduction into account.

As we will see in the following, some extracted features play a much bigger role than the other ones for an accurate weight estimation. For this reason, we are interested in testing several estimators as the extended linear regression which is used in (Kirushanth, 2020) and the local linear regression (Llr) (Cleveland, 1988), that has proven to withstand the curse of dimensionality and widen the class of functions that can be approximated. However, (Cleveland, 1988) shows accurate forecasts in the case of two or three independent features. We will show that performing the principal components analysis (PCA) before the local linear regression outperforms other classical parametric models such as extended linear regression.

The present paper is organized as follows. The first part is the reminder of the mathematical principle which will be used to create our procedure. It consists of the definition of a smoother and its application in the local linear regression, the basics of the principal components analysis (PCA) and the statement of the relative error enforced by OIML. Then, the available data sets and their specifications will be presented. Third, we will compare the learning curves of the parametric models and the Llr. Finally, we discuss the results and the time and space complexity of the method.

2. Mathematical Background

2.1. Model and database

Let **Y** be the vector in $\mathbb{R}^{Nx1}$ of *N* reference weights and **X** the $\mathbb{R}^{NxD}$ measurement matrix of *D* features returned by the HS-WIM. Each column of this matrix refers to a typical feature, such as the speed or the estimated weight, for example.
Y is assumed to be known thanks to a training phase with regard to the respective measurement matrix **X**. We call training phase the period in which we record the measures of the HS-WIM and the reference weight of the LS-WIM.

Suppose that the data are generated by the model :

$$y_n = f(x_n) + \epsilon_n \tag{1}$$

In which f, x_n *and* ε_n are respectively a general smooth non-linear function as explained in (Cleveland, 1988), all the measured features of the nth weighed vehicle and the random error of measurement with mean $\mu = 0$.

The regression problem is to find the conditional expectation $\hat{y}$ of the true weight *y,* given the new measurement *x* for a typical vehicle, such that :

$$\hat{y} = E[y|x] = F(x) \tag{2}$$

Where F stands for the used model.

2.2. Principal components analysis (PCA)

The first step of our regression problem is solving the problem of multi-colinearity between the available features of a vehicle, which are mainly the lateral position of each wheels $\{P_j\}$ on the reference lines and the speed of the vehicle v. Though the temperature of the road T has been proven in (Gadja, 2016) and (Burnos, 2016) to be a significant factor of influence in the weight estimation, our collected data does not contain a wide range of temperatures. Then, the optimized parameters in each model could only give a poor estimation of the real parameters. Therefore, the temperature will not be considered in what follows, even if such algorithms presented in (Burnos, 2020) could be implemented in addition to our method.
We use the PCA to extract new uncorrelated features from the initial features, as explained in (Bishop, 2006).

If **X** is our matrix of measurements, we can see each sample as a vector :

$$x_n = \left[x_n^1 \, x_n^2 \ldots x_n^D \right] \tag{3}$$

The PCA seeks to find a new representation of the same dataset in which the new features are uncorrelated. Let **Z** be this new data matrix such that we compute the new variables as linear combinations of the initial features:

$$\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{V} \tag{4}$$

Where $\boldsymbol{V}$ is the stacked column vectors for the linear transformation.
This matrix is computed with the Eigen Value Decomposition (EVD) applied to the correlation matrix $\boldsymbol{\Sigma_X}$ of **X** such that:

$$\boldsymbol{\Sigma_X} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V^T} \tag{5}$$

with :

$$\boldsymbol{\Sigma_Z} = \boldsymbol{\Lambda} = diag\left(\langle \lambda_i \rangle_{i=1}^{D} \right) \tag{6}$$

Therefore, the new correlation matrix $\boldsymbol{\Sigma_z}$ (if we assume that the data has been standardized) is equivalent to $\boldsymbol{\Lambda}$ which is the diagonal matrix of the variances of the new features.
After the feature extraction, we ought to select the new features set $\{z_d\}$ which optimizes the accuracy in cross-validation.

2.3. Local linear regression (Llr)

In this part, we introduce the model *F* that we choose to estimate the true weight from the features measured by the HS-WIM. The local linear regression has some key advantages compared to other parametric method as linear regression, which are explained in the following.
Firstly, this method is completely adaptive because of its non-parametric optimization. This means that any additional step is required to implement the method in a different location.
Secondly, the Llr has the ability to widen the function to be approximated. Then, applying non-linear function and the stepwise selection as presented in (Kirushanth, 2020) becomes unnecessary. This results in an end-user program easier to handle which does not need to be validated. Instead of computing the weight regarding the features of the vehicle as it is the case for classical linear regression, the main idea of the local regression is predicting the weight of a vehicle with previously encountered vehicles during the training phase which are "similar" with the new vehicle. This similarity is computed as follows, with the function *w*.

The local linear regression was initially defined with the principle of smoothers introduced in (Cleveland, 1988) among others in order to smooth time series.
The local linear regression is based on a scalar product between the vector of β and the vector of features, extended by the coefficient 1 :

$$F(x|X,Y,w,k)=[1\,x]\,\beta(x|X,Y,w,k) \tag{7}$$

In our case, *k* is simply the number of neighbors that are considered for the estimated weight $\hat{y}$ of measurement *x*. As mentioned in (Cleveland, 1988), a possible smoother is the tri-cube function :

$$w(x,x_n|k)=max\left[0,\left(1-\left(\frac{\|x-x_n\|}{d_{max}^{k+1}}\right)^3\right)^3\right] \tag{8}$$

The model ought to optimize the value of *b* for the new measurement *x*, given **X** and **Y** obtained during the training phase, and given a weighting function of neighborhood *w* with parameter *k*. d^{k+1}_{max} is the $(k+1)^{th}$ smaller distance between x and the set of measurements $\{x_n\}$.

$$\beta(x|X,Y,w,k)=argmin_b\sum_{n=1}^{N} w(x,x_n|k)\left(y_n-([1\,x_n].b)\right)^2 \tag{9}$$

It can be shown that this minimization problem is equivalent to solving the following linear system in which the variables are the parameters in the vector β:

$$[\widetilde{X}'A\widetilde{X}]\beta = A\widetilde{X}'Y \tag{10}$$

Where the matrix tilde **X** is the extended matrix **X** by the vector of ones $\mathbf{1}^N$. **A** is the diagonal of the weight between each x_n and x :

$$A = diag\left(\left\langle w(x, x_n)\right\rangle_{n=1}^{N}\right) \tag{11}$$

Obviously, only k in N equations are used in the linear system due to the weight function w that is equal to zero for the samples x_n for which $\| x - x_n \|_2 > d^{k+1}_{max}$.

Some additional remarks can be made about the hyper-parameter *k*. The key point to be understood is that *k* is not related to the model itself, as the classic parameters do in a linear regression, but mainly to the number of vehicles we recorded during the training phase. While parameters are not consistent in different HS-WIM system, the hyper-parameters can be fixed if they have a similar number of recorded vehicles.

2.4. Definition of relative error and accuracy

The fundamentals for the assessment of the accuracy are introduced according to the COST323 and (Jacob, 2000).
Let y be the reference weight and $\hat{y}$ the estimated weight computed with a mathematical model, which is possibly inaccurate. For each weighed vehicle, we can compute its relative error of estimation e^r :

$$e^r = \frac{\hat{y} - y}{y} \tag{12}$$

Then, the accuracy of a HS-WIM machine is defined as the probability Π that a single relative error falls inside the range [-δ,+δ]. More precisely, we seek to find a lower bound π of Π with risk α, as the true mean of this distribution is unknown.
According to (Jacob, 2000), it is possible to find a lower bound π computed as follows :

$$\Pi \geq \pi = \Psi\left(\frac{\delta - m}{s} - \frac{t_{\nu,1-\alpha/2}}{\sqrt{N}}\right) - \Psi\left(\frac{-\delta - m}{s} - \frac{t_{\nu,1-\alpha/2}}{\sqrt{N}}\right) \tag{13}$$

m and s are the estimated mean and standard deviation, respectively; $t_{\nu, 1-\alpha/2}$ and Ψ are the (1-α/2)-quantile for the Student's t distribution with ν = N-1 degrees of freedom and the Student's t CDF with ν degrees of freedom, respectively.

According to the specifications of the COST323, we especially aim to achieve the class A(δ = 5%) for the T2S3 and class B(δ = 10%) for the U2, with π = 99.8% and α = 5%.

We emphasize the need of checking the normality of the relative errors before computing the confidence levels. To this end, we are going to perform a Shapiro-Wilk test on each set of relative errors in the next sections.

2.5. Subset of de-correlated variables for neighborhood

As it was already mentioned, (Cleveland, 1988) shows good estimations for a small number of independent features. Though, we explained that we have a much bigger number of variables that are not independent. At first sight, we might think that proceeding the PCA on all the dependent features and using the *d* ones with the higher variances in the smooth function could be a good idea, but it is not the case because increasing the variance between the features does not necessarily lead to an increasing of the correlation between these features and the reference weight.

Therefore, two consequences follow. Firstly, considering useless features in the distance function *w* gives less importance to the real significant features. Secondly, it would be better to apply separately the linear tools PCA on the subsets $\{W_i\}$ and $\{P_j\}$ and optimizing the coefficients of this relationship with regard to the reference weights $\{ y_n \}$.

Therefore, there are two dimensionality reductions: d_1 and d_2, which are the dimensional reduction of the subsets $\{W_i\}$ and $\{P_j\}$, respectively.

To this end, let κ_d denote the aggregated variance kept from the first *d* new features, compared to the total variance:

$$\kappa_d = 100\,\% \times \frac{\sum_{l=1}^{d} \lambda_l}{\sum_{l=1}^{D} \lambda_l} \qquad (14)$$

We can see the kept variance regarding the dimension reduction *d* for each subset in the figure 2. We notice that the subset of $\{P_j\}$ of 10 variables have actually a high dependency and that only this transformed feature alone explains more than 95% of the total variance on $\{P_j\}$.

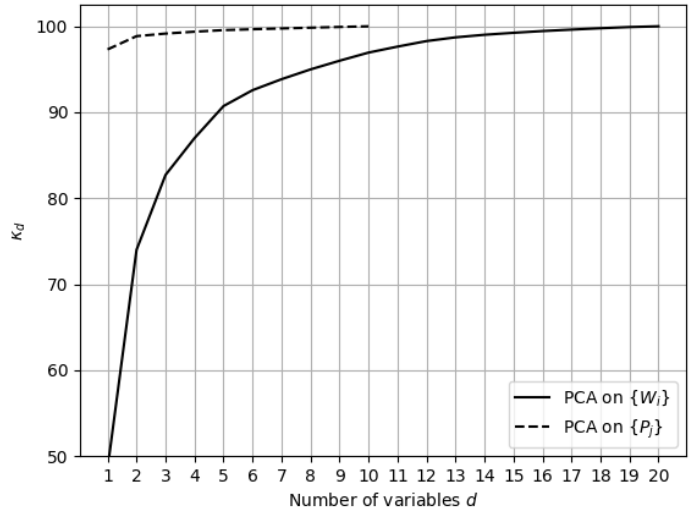

Figure 2 – PCA on the subsets of the 20 weights under each wheels and of the 10 wheel positions

If we look at the PCA coefficients for the subset $\{P_j\}$, we noticed that this principal component is the mean of the position of the wheels, $z^P{}_1$. Now, we can compare the error of estimation regarding this variable in the figure 3.

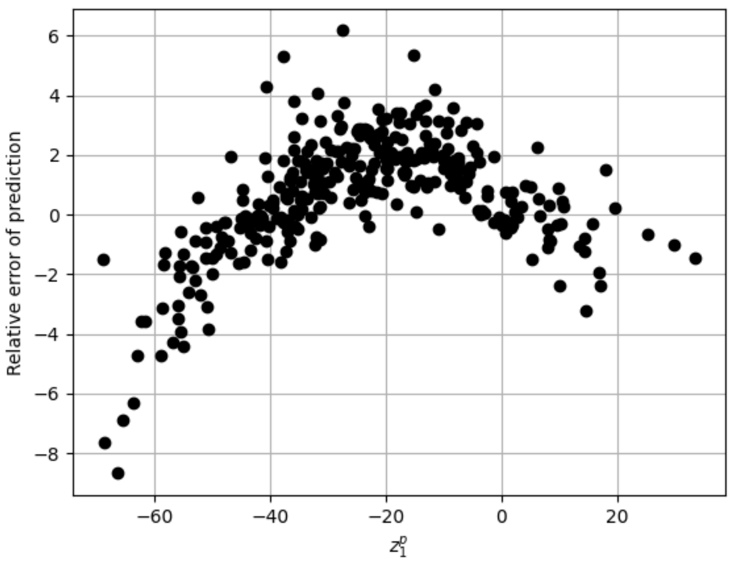

Figure 3 – Error of estimation (%) compared to the first principal component of the subset {P_j}

In this figure, we observe that the uncertainty of the error, given the principal component z^P_1, dramatically decreases. We can also notice that this curve is parabola-like but it can be more complex according to the type of vehicles. While the auto-covariance matrix of **X** gives us information about how the features are related to each other, it does not ensure that the PCA variables depends on the target value *y*. To this end, we compute the covariance between each z^w_i and *y*. The absolute value of these quantities are shown in the figures 4 and 5 for both the T2S3 and U2, respectively. We can observe that the covariance only highlights linear dependencies of few PCA variables with *y*. For T2S3 and U2, we observe that $d_1 = 3$ explained a great part of *y*.
This observation will be compared with the K-Fold cross-validation in the section 5.

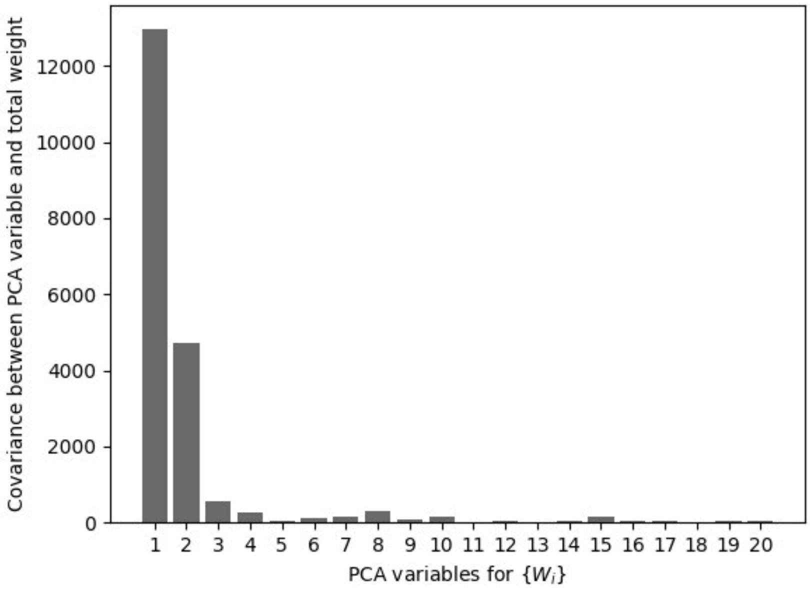

Figure 4 – Absolute value of covariance between each variable z^W_i and *y* for T2S3

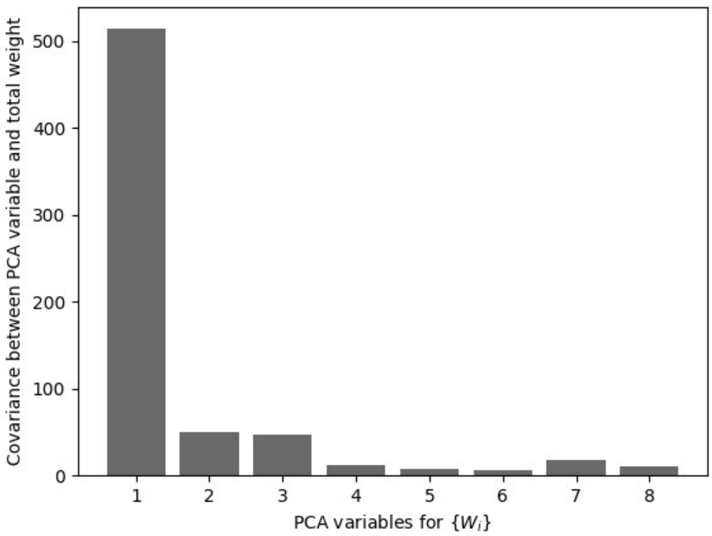

Figure 5 – Absolute value of covariance between each variable z^W_i and *y* for U2

For these reasons, we will try a first model inspired from the local regression, but in which we only use the variable z^P_1 in the distance function. So, let $\{z^{W}_{i}\}$ be the PCA variables of the subsets $\{W_i\}$. We are now able to simplify the definition of the local regression as:

$$F\left(\left\langle z_i^W\right\rangle, z_1^P, v\middle|, Z, Y, k\right)=\left[1\left\langle z_i^W\right\rangle z_1^P v\right].\beta\left(z_1^P\middle|Z, Y, w, k\right) \tag{15}$$

We only use the variable z^P_1 for the computation of the neighborhood and afterwards we use a linear combination of the other variables. This is because the non-linear relationship between the variables $\{z^W_i\}$ and the target *y* mostly depends on the value of $\{z^P_j\}$.

3. Data collections

In this section we present the data we used to train and test our model. We assess the accuracy of our model for the gross-weight (GW) of two kinds of vehicles; the trucks with a header of two singles axles in addition to a trailer that is a group of three axles (T2S3) and vans with only two single axles (U2). Their maximum permissible weight is usually around 44t and 3.5t respectively for the T2S3 and the U2.

The HS-WIM system contains two grids of sensors which independently weigh a vehicle. Each grid computes the raw estimation of the weight with the mean of all the single weights measured by the sensors. The double weighing by the grids is a legal requirement for weighing station in order to prevent outliers and abusive penalties.

In the following results, we are only trying to model the grid 1. The use of only one grid is important because a future research could seek to predict all types of vehicles with the same

model, and then from the same grid. The following procedure can obviously be used for the second grid.
In what follows, the reference weight is approximated with the weight given by the LS-WIM since the error given by this scale is small (1%) compared to the error of the HS-WIM.
The specifications of the low speed scale are depicted in the table 1.

Table 1 – Raw accuracy of database 1 for the types of vehicles T2S3 and U2

Effective speed	5 km/h
Max	20t
Min	20kg
d	10kg
Class	III
Axle max error	4%
Total max error	1%

A massive campaign was achieved from September 2019 to February 2020 to collect data. The raw accuracy of the system is summarized in the table 2. During this period, vehicles of concern were deviated from the road and weighed with a high-precision machine (LS-WIM) after their weighing with the HS-WIM.

Table 2 – Raw accuracy of database 1 for the types of vehicles T2S3 and U2

	GW-T2S3	GW-U2
Number	404	188
m (%)	-0.18	-4.71
s (%)	5.08	2.90

The table 3 depicts another database in which the measurements were collected in June and July 2020, in which all measurements do not come from different vehicles, but from a few chosen vehicles that passed on the sensors. This database will be interesting because of its simplicity to carry out. Moreover, it provides us a cheaper training with a lower bound on the accuracy. Since the table 3 will only be used as a training data, we will use as much information as possible. Therefore, if we look at the confidence level of both types without validity state in both tables, we see that the accuracy could be much better in a table compared to the other one. That is explained by the high dependency between the behavior of a single vehicle and the pattern of its relative errors. In the table 2, we see that the few tested T2S3 give good results, while the few tested U2 give bad accuracy.

Table 3 – Raw accuracy of database 2 for the types of vehicles T2S3 and U2

	GW-T2S3	GW-U2
Number of measurements	167	144
Number of vehicles	2	2
Number of loads	8	6
m (%)	0.85	-6.9
s (%)	1.69	8.31
P-value Shapiro test	4	0

In the table 4, we present the accuracy of the raw system for the massive campaign DB1, after the application of the statistical tools and validity criterion of the original system. This validity criterion is based on conditions which have to be met in order to pass the selection. We observe that Class A for the T2S3 and the class B for the U2 are obtained in this way. Then, we notice that high accuracy is obtained, but only the half of the trucks has a valid weighing.

Table 4 – Raw accuracy of massive campaign with hard validity conditions for the types of vehicles T2S3 and U2

	GW-T2S3	GW-U2
Number	404	188
Valid ones	180	185
m (%)	0.53	-2.41
s (%)	0.98	2.04
δ (%)	5	10
Π_δ (%)	99.99	99.98

Besides the confidence levels and intervals, another more practical criterion for the good fit of the estimation is the confusion matrix for the classification of penalties. Theses matrices are shown in tables 5 and 6 for the two types of vehicles. These tables show that a small part of the vehicles, which have their weight beyond the limit, are given a penalty at the end. This is because only a few part of the vehicles is considered as valid, and then weighed. However, we can notice that the raw estimation never gives false positive which is a necessary condition for the system deployment.

Table 5 – Confusion matrix for valid T2S3 of the database 1

		Estimation penalty		
		Yes	No	Total
Real Penalty	Yes	16	35	51
	No	0	353	353
	Total	16	388	404

Table 6 – Confusion matrix for U2 of the database 1

		Estimation penalty		
		Yes	No	Total
Real Penalty	Yes	72	48	120
	No	0	68	68
	Total	72	116	188

4. Computational complexity

Another significant consideration is the ability of the model to run in small electronic devices. Then, we analyze in this section the time and the memory space required to run our method.

Let N ,k, d_w, d_P be the number of measurements available during the training phase, the number of considered neighbors as introduced in the previous sections, the dimension of the reduction for the sets $\{W_i\}$ and $\{P_j\}$, respectively. For the computation of the complexities, there is no need to split d_w and d_P because we mainly used the total number of dimensions, that we will call d, which is define as $d = d_w + d_P + d'$ where d' is the number of dimensions that does not belong to one of the both PCA subsets, such as the speed for instance. The only small difference in the terms of complexities is the computation of the distances, which only depends on the dimension d_P.
The time and memory complexity of the non-linear regression is analyzed as follows.

Concerning the time consumption, we have two significant steps in the modeling: the search of neighborhood and the minimization problem (9). For the first step, we ought to compare all the distances between the new measurement and each measurement from the training database. Sorting can be achieved by the algorithm *Quicksort* (Hoare, 1962) which has proven to sort in order $O(N \log N)$ for the best and average case, and $O(N^2)$ for the worst case.
However, we have in our case $k << N$, so we had better look at k times the minimum of the distances list. We can reduce the time complexity of this computation to $O\ (k \times N)$ in the average and worst case.
The second step is the solving of the minimization problem. The computation time mainly comes from the several matrix multiplications and the solving of the linear system $C\,\beta = b$.

Such linear system with Hermitian matrix $C \in \mathbb{R}^{d \times d}$ can be solved by a Cholesky decomposition as explained in (Burnian, 2004). This decomposition is $O(d^3)$ and the calculation is $O(d^2)$.
At the end, the total time complexity is $O(k \times N + d^3)$ in the worst case, as we typically have $d << N$.
In the next section, we are going to see that $d = 3$ already outperforms other parametric estimators. However, we must pay attention to the number of considered data points N which can be possibly large, more than 100 for instance.
Concerning the memory complexity, the critical point is saving a possibly large database for the training, so the order of space complexity is $O(N \times d)$. For instance, we should possibly store 400×3 float numbers in order to predict new weights. The very small hard memory needed for this storage should be at least equivalent to 5 KB.

5. Results

In this section, we are addressing the problem of comparing the accuracy of the local linear regression applied with different hyper-parameters and other parametric estimators. We especially pay attention to the definition of accuracy in this context because of the difficulties to collect true weights on a static machine. Therefore, accuracy is considered as learning curves, which are the mean of accuracy with different subsets of training vehicles, with regard to the size of these sets. The experiments are performed as follows.

We define a number of experiments $M = 100$. For each $m < M$, we apply a mixing function to the database $\{ x_n \}$. For each m, we also define a train-test split of the database, with a balance of 80% /20%. Then, for each m we have a split X^{test}_m and X^{train}_m.
However, in order to assess the accuracy of the estimator regarding the available number of training vehicles, we let the model discover the training set step by step. At Each step l we adds 10% of the total number of training vehicles. At the end, we have in fact M sets of estimations for each l:

$$\left\langle \hat{y}^{test}_{ml} \right\rangle = F\left(X^{test}_{m} \middle| X^{train}_{ml}, Y^{train}_{ml}, w, k \right) \tag{16}$$

With the definition of relative errors, we are able to compute the accuracy π_{ml}^{test} for the set of errors. At the end, we have M learning curves of L evaluations of accuracy. For each step l, we compute the mean and the 5%-quantile of the M sets.

In the figures 6 and 7, we observe the learning curves for several values of k, d_1 and d_2 and different contexts of application of the PCA. For different values of k, three different local linear regressions are performed. The first one Llr-PCA is the Llr with the locality which is only computed as the distance in the variable z_1^P.
In addition, the total weight and the speed are used as parametric variables. Therefore, there is an application of PCA for the subset $\{P_j\}$ but not for $\{W_i\}$. The three other ones are the Llr with application of PCA for the subsets $\{W_i\}$ and $\{P_j\}$, in which the number of kept dimensions are d_1 and d_2 respectively. We notice that the worst mean accuracy is achieved by the value of $k = 20$ for the two types of vehicles and for all the cases of hyper-parameters. This is because the number of neighbors is too small to be able to give an accurate mean behavior of the neighbors. We notice that the mean accuracy of $k = 50$ and $k = 80$ are quite similar for a lot of training vehicles. The reason is the local behavior is not conserved if we

keep too different vehicles for the training. It can be easily understood that there must be a trade-off for the value of *k*. If the value is too high, we loose the local characteristics of the most similar vehicles.
On the contrary, we loose the notion of mean and the model can be too sensitive with regard to some specific vehicles. The value of *k = 50* appears to be in the optimal range for both types of vehicles.

Concerning the best application of PCA, we see that the application of PCA on the subset $\{W_i\}$ has no benefit in this case and even tends to slightly decrease the mean accuracy. The model Llr-PCA seems to be the more accurate for a small number of training vehicles.

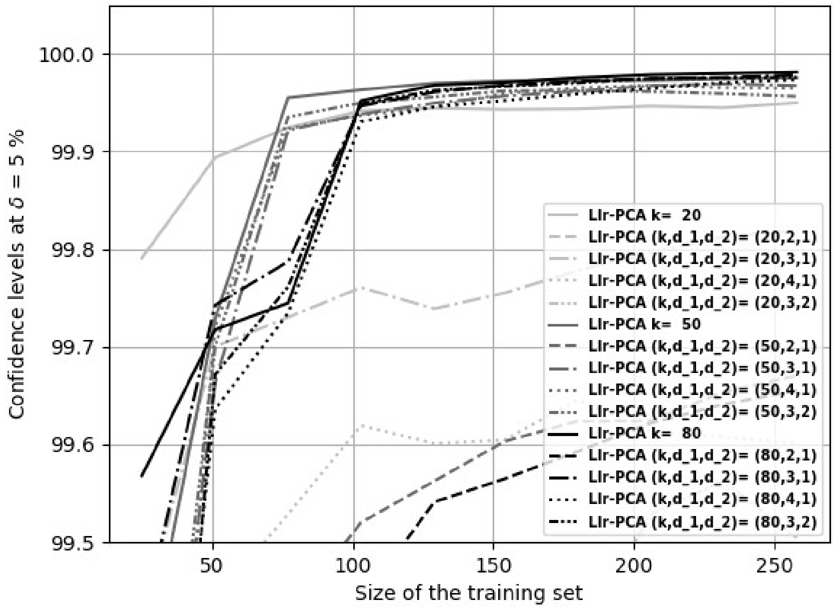

Figure 6 – Comparison of accuracy for several local linear regressions for different values of k for T2S3

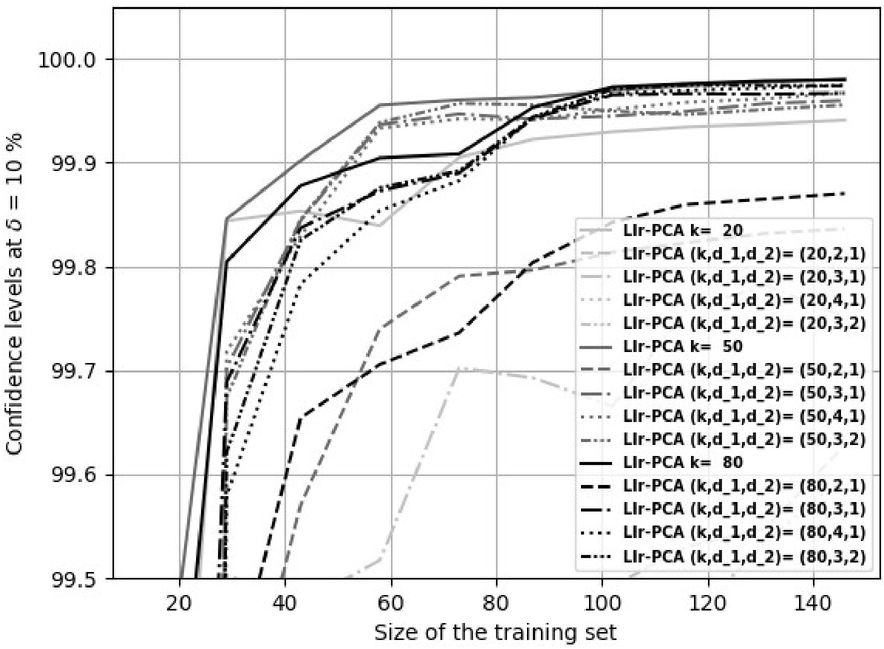

Figure 7 – Comparison of accuracy for several local linear regressions for different values of k for U2

Afterwards, we will compare the accuracy of the best estimators in the previous figures 6 and 7 with regard to other parametric models such as principal components regression and extended linear regression.

We define a second estimator LR-5 that takes into consideration the gross-weight, the speed and the variable z^P_1 to the power 1 to 5 which is added to the model, in order to model the non-linearity with parametric regression. A more general model as principal components regression can be performed to the extended matrix of all features. The last estimator Llr is an implementation of (Cleveland, 1988) for our standardized high-dimensional data without applying PCA.

In the figures 8 and 9, we see the mean of the *M* learning curves at each points *l* for the two types of vehicles. We notice that the accuracy of the three estimators Llr-PCA's and the LR-5 are much better that the other parametric methods.
$\pi = 99.8\%$ is reached at the end of the training but this good accuracy is obtained quickly, with less than 100 vehicles for the training phase of the both types. For the other tested methods, we see that, with enough components, the PCR is able to reach good accuracy, but for a very large number of training vehicles.

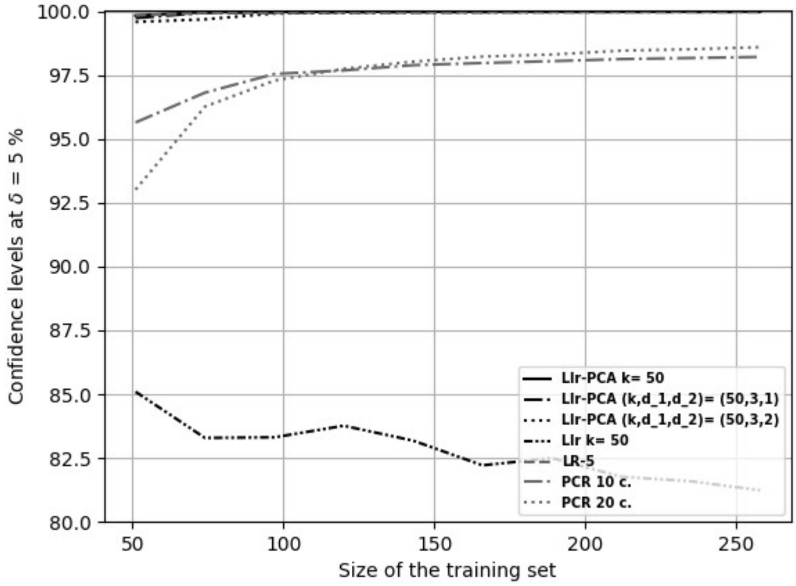

Figure 8 – Learning curves of several models with a split Train-Test of the database 1 for T2S3

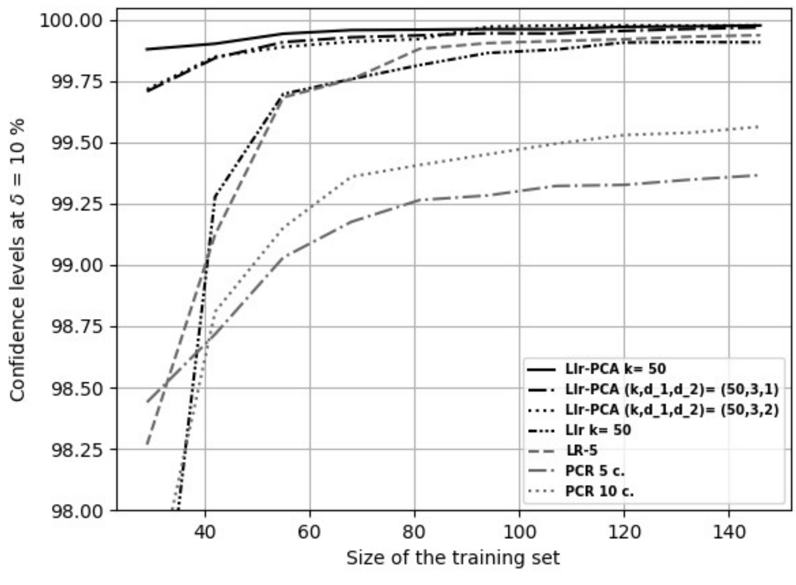

Figure 9 – Learning curves of several models with a split Train-Test of the database 1 for U2

In the figures 10 and 11, we show the mean accuracy of the four best estimators in the above-mentioned figures and their 5%-quantile. We notice that the mean of the accuracy of the Llr-PCA is the highest one. Moreover, we can also observe that the 5%-quantile is beyond the confidence levels 99.8% and closely follows the mean. In particular, we observe that the mean and the 5%-quantile are both above 99.8% for 100 training vehicles.

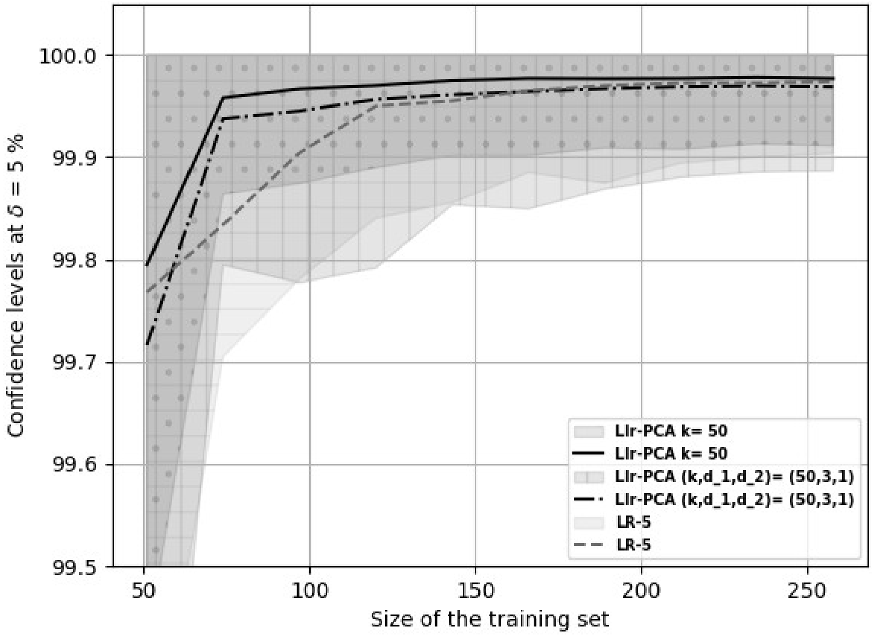

Figure 10 – Learning curves of several models with a split Train-Test of the database 1 for T2S3

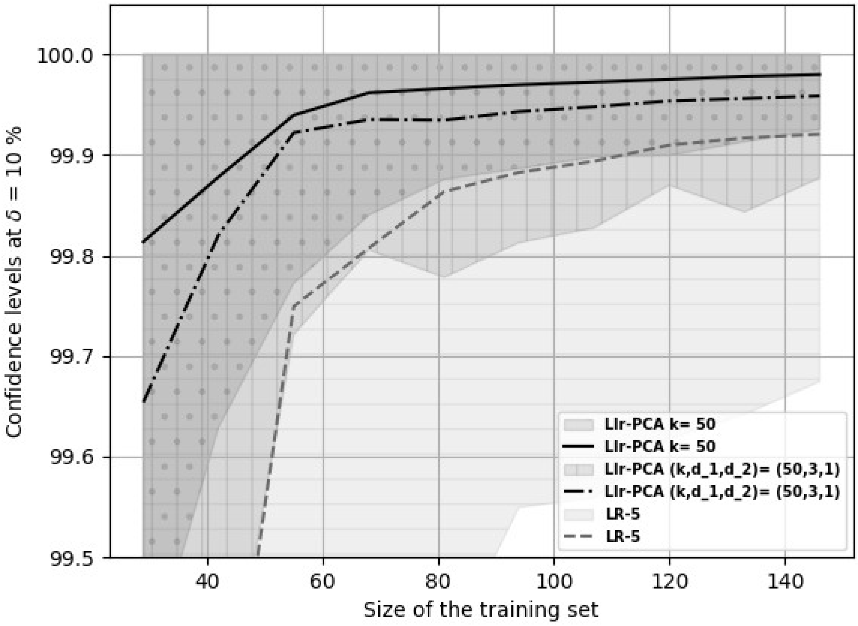

Figure 11 – Learning curves of several models with a split Train-Test of the database 1 for U2

The table 7 depicts the accuracy of the Llr-PCA in *Leave-One-Out (LOO)* experiment for the database 1. If we compare it with the accuracy of the raw system in the table 3, we notice that the mean of the errors is highly reduced while the standard deviation of the errors is slightly higher for the T2S3 and for the U2. Moreover, we are able to perform better estimations for 323 trucks instead of 180 for the raw system. However, we have to pay attention that the H_0 hypothesis Shapiro-Wilk test returns "false" for the trucks. Therefore, the confidence levels must be carefully analyzed.

Table 7 – Accuracy in Leave-One-Out with the local linear regression of database 1 for the types of vehicles T2S3 and U2

	GW-T2S3	GW-U2
Number	323	183
m (%)	0.02	0.11
s (%)	1.13	2.12
P-value Shapiro test	0	74
δ (%)	5	10
Π_δ (%)	99.99*	99.99

To this end, the figures 12 and 13 show the relative errors of the Llr-PCA compared to the original errors and their respective Gaussian approximation. For the T2S3, we observe that the relative errors are more localized around the zero error than the Gaussian approximation. Therefore, we can expect that even if the Shapiro-Wilk test is negative, the real performance would be as good as the claimed confidence level in the table 7. Moreover, we notice that the total number of vehicles is much bigger for the Llr-PCA while the accuracy is still as high as the previous one.
For the U2, we notice that the number of tested vehicles remains the same but there is a real improvement for mean of the errors.

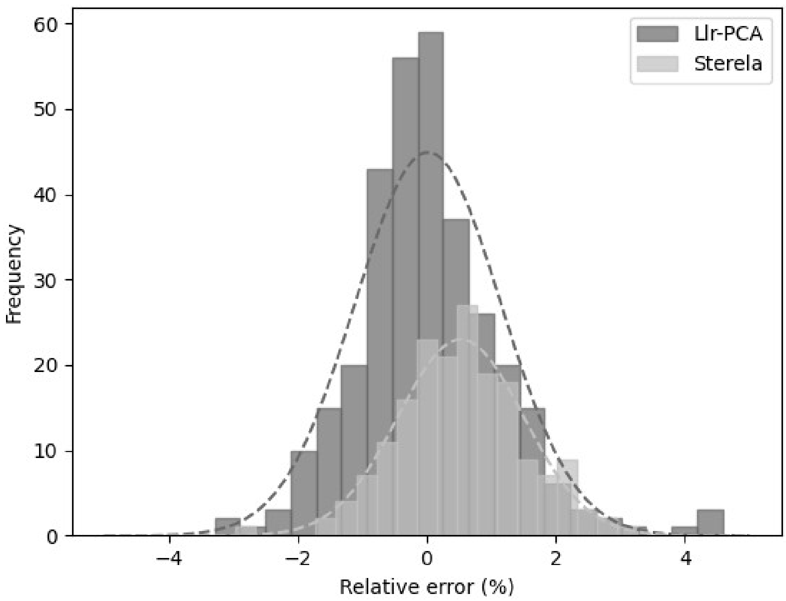

Figure 12 – Comparision of the estimation errors between the raw system and the Llr-PCA for T2S3

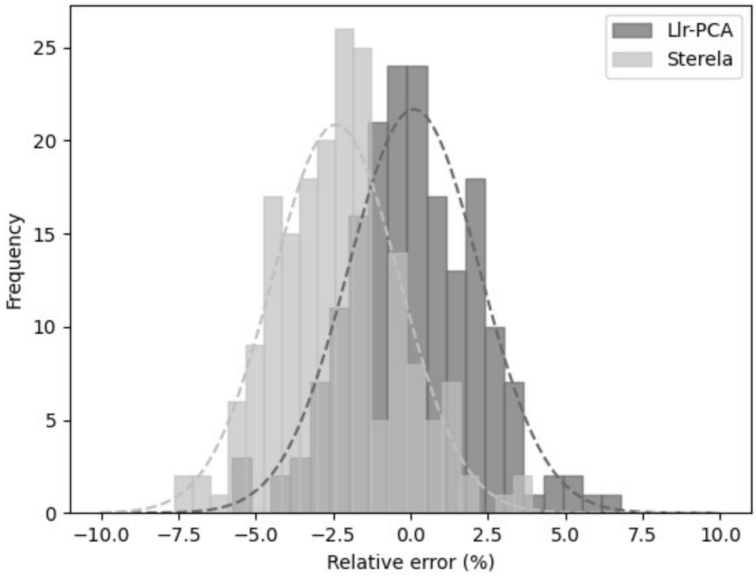

Figure 13 – Comparision of the estimation errors between the raw system and the Llr-PCA for U2

In the tables 8 and 9, we notice that the given penalties for both types of vehicles, in the case of training by LOO. We notice a significant improvement with regard to the initial confusion matrices in 5 and 6.

Table 8 – Confusion matrix of the local linear regression in LOO with δ = 5% for the T2S3 of the database 1

		Estimation penalty		
		Yes	No	Total
Real Penalty	Yes	28	23	51
	No	0	353	353
	Total	28	376	404

Table 9 – Confusion matrix of the local linear regression in LOO with δ = 10% for the U2 of the database 1

		Estimation penalty		
		Yes	No	Total
Real Penalty	Yes	83	37	120
	No	0	68	68
	Total	83	105	188

Secondly, another experiment can be performed. We foresee that the following training is much less suitable than the first one because we train the models with a permutation of the measurements of the database 2 and then we evaluate the accuracy with the database 1.
Therefore, a small part of the features space of the database 1 will be considered during the training phase. We can define the same procedure as the equation (16), with $\mathbf{X}^{train}$ which stands for a mixture of the database 2 and $\mathbf{X}^{test}$ which stands for the whole database 1.
We can observe the learning curves of these experiments in the figures 14 and 15. As already said, it is not surprising to observe a decrease of the accuracy. In this context, the local linear regression seems to stay the more robust model against a poorer training.

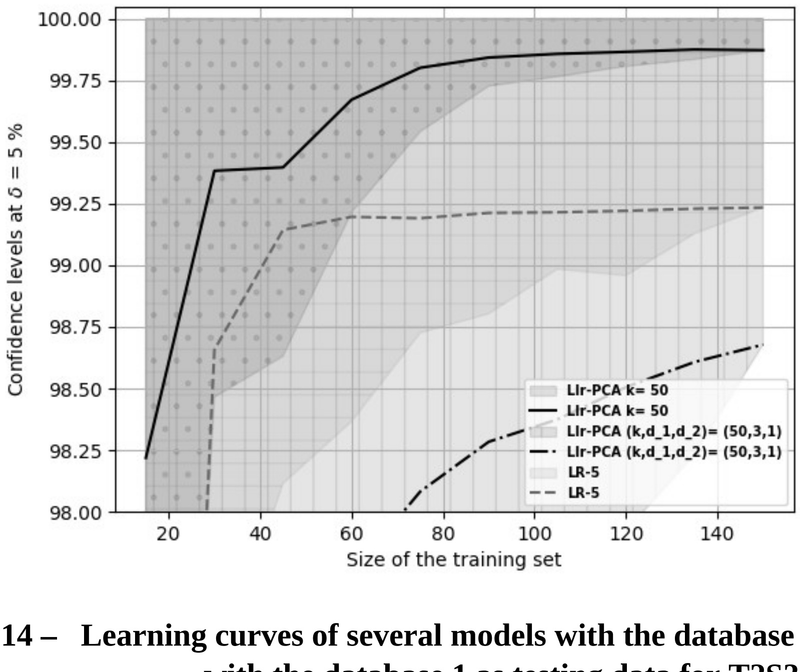

Figure 14 – Learning curves of several models with the database 2 as training data and with the database 1 as testing data for T2S3

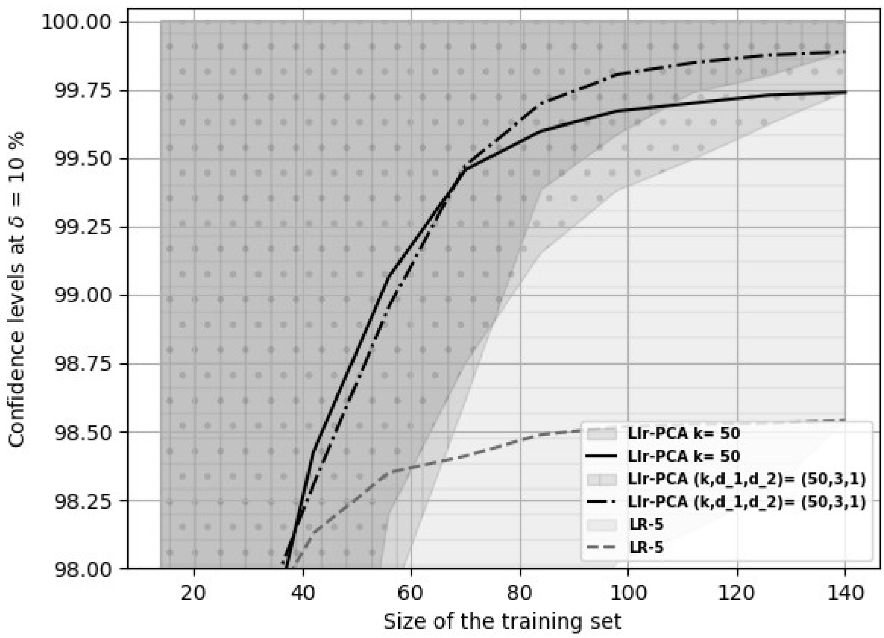

Figure 15 – Learning curves of several models with the database 2 as trainingdata and with the database 1 as testing data for U2

The table 10 shows the accuracy with the second training, for each type. We observe that the accuracy are still above 99.8% for the T2S3.

Table 10 – Accuracy with the database 2 as training data with the local linear regression of database 1 for the types of vehicles T2S3 and U2

	GW-T2S3	GW-U2
Number	323	183
m (%)	0.21	1.23
s (%)	1.48	2.93
P-value Shapiro test	1	51
δ (%)	5	10
Π_δ (%)	99.87*	99.74

6. Conclusions

Weight controlling on the highway can be a very painful work with a classical low-speed weigh-in-motion. The high quality of the piezoelectric quartz sensors and the recent re-foundation of the roadway provide a very accurate weight estimation, i.e. an error of estimation that has a probability of 99.8% to be inside the range [-5 %, 5%] for the trucks and

[-10%, 10%] for the U2. However, the main drawback was the very restrictive conditions that a vehicle must meet to be weighed. As a consequence, only the half of the total number of trucks can be weighed.

Therefore, we proposed a comparison between several estimators and we concluded that the non-parametric function gives the best accuracy with regard to the mean and the 5%-quantile, for two different types of vehicles and for two types of training: cross-validation with LOO and validation with an artificial training. Some additional advantage of this method can be highlighted, as the capability of the model to be easily used in another HS-WIM system without additional validation steps.

Our method increases the confidence levels in all cases for a much greater number of vehicles, and especially to more than 99.99 % for the two types of vehicles in *Leave-One-Out*. Thanks to the cancellation of the validity state and to the introduction of the local linear regression, we are able to increase the number of vehicles that are weighed, from 180 with the selection to 323 for the T2S3 with our model. Both the improvement of the accuracy and the bigger number of tested vehicles lead to an increasing of the given penalties from 16 to 28 and from 72 to 83 for the T2S3 and the U2, respectively. Moreover, we observed that the Llr is able to reach this accuracy with 100 vehicles.

Besides these good results, some key points remains interesting to investigate.
The main one is the ability of the model to forecast the weight of other kinds of vehicles. In selecting vehicles of concern, their similar physical structure can be easily approximated by a single PCA variable. In allowing all types of vehicles in the same procedure, we can foresee that more features could be necessary in order to keep good performance. For instance, expanding the concept of local linear regression to an additive model of several local smoothers, as in a projection pursuit regression (Friedman, 1988), could be a good idea in order to take more components into consideration.

6. Acknowledgment

Firstly, we would like to thank Adriana Antofie and Dominique Corbaye for their expertise in the field of WIM and their daily work in the SPWMI.
Secondly, we would like gratefully acknowledge Anne-Sophie Libert, Timoteo Carletti and Germain Van Bever, Professors at the University of Namur for their general involvement in this project.
Finally, we also thank the company Sterela for sharing their data and expertise during this partnership.

7. References

- Antofie A., Boreux J. and Corbaye D. (2019), “Approach of the walloon legal metrology (belgium) for weigh-in-motion (wim) free flow direct enforcement”, 8th International Conference on Weigh-in-Motion.
- ASTM International (2021), “Standard specification for highway weigh-in-motion (wim) systems with user requirements and test method”, Annual Book of ASTM Standards, Vol 04.07.
- Bishop C. (2006), “Pattern Recognition and Machine Learning”, Springer.

- Burian A., Takala J. and Ylinen M. (2004), “A fixed-point implementation of matrix inversion using Cholesky decomposition”, Midwest Symposium on Circuits and Systems 3 1431–1434.
- Burnos P. and Gadja J. (2016), “Thermal property analysis of axle load sensors for weighing vehicles in weigh-in-motion systems”, MDPI Sensors.
- Burnos P. and Rys D. (2017), “The effect of flexible pavement mechanics on the accuracy of axle load sensors in vehicle weigh-in-motion systems”, MDPI Sensors.
- Burnos P., Gadja J., Sroka R., Wasilewska M. and Dolega C.(2021), “High accuracy weigh-in-motion systems for direct enforcement”, MDPI sensors.
- Burnos P. and Gadja J. (2020), “Optimised autocalibration algorithm of weigh-in-motion systems for direct mass enforcement”, MDPI sensors.
- Cebon D. and Winkler C. (1991), “Multiple-sensor weigh-in-motion: Theory and experiments”, Transportation Research Record.
- Cleveland W.S. and Devlin S.J. (1988), “Locally weighted regression: An approach to regression analysis by local fitting”, Journal of the American Statistical Association 83 596–610.
- Doupal E., Adameova Z. and Kriz I. (2016), “Start of direct enforcement in the czech republic”, ICWIM7 7th International Conference on Weigh-in-Motion.
- Friedman J.H. and Stuetzle W. (1981), ”Principal pursuit regression”, Journal of the American Statistical Association 76 817–823.
- Gadja J., Burnos P., Sroka R. and Daniol M. (2023), “Accuracy maps of weigh-in-motion systems for direct enforcement”, MDPI Electronics.
- Gadja J. and Burnos P. (2016), “Temperature properties of weigh-in-motion systems”, ICWIM7 7th International Conference on Weigh-in-Motion.
- Gadja J., Sroka R. and Burnos P. (2020), “Sensors data fusion in multi-sensors weigh-in-motion systems”, MDPI sensors.
- Heidari F. and Pouyan M. (2019), “Application of deep learning technique in high-speed weigh-in-motion systems for direct enforcement”, ICWIM8: 8th International Conference On Weight-In-Motion.
- Hoare C. (1962), “Quicksort”, The Computer Journal 5 10–16.
- Jacob B., O’Brien E. and Newton W. (2000), “Assessment of the accuracy and classification of weigh-in motion systems: Statistical background”.
- Jacob B., O’Brien E. and Newton W. (2000), “Assessment of the accuracy and classification of weigh-in motion systems: Part 2 european specification”.
- Jacob B. and Cottineau L.M., (2016), “Weigh-in-motion for direct enforcement of overloaded commercial vehicles”, Transportation Research Procedia.
- Kirushanth S. and Kabaso B. (2020), “Design and development of weigh-in-motion using vehicular telematics”, Journal of Sensors.
- Lee T. and Solo V. (1999), “Bandwidth selection for local linear regression: A simulation study”, Computational Statistics 14 515–532.
- Marchadour Y. and Jacob B. (2008), “Development and implementation of a wim network for enforcement in France”, Proceedings of the 5th International conference on WIM.
- O’Brien E. and Jehaes S.(1998), “Cost 323 weigh-in-motion of road vehicles”.
- Savitzky A. and Golay M.J. (1964), “Smoothing and differentiation of data by simplified least squares procedures”, Analytical Chemistery 36 1627–1639.
- Stergioulas L., Macleod M. and Cebon D.(2000), “Static weight estimation and system design for multiple-sensor weight-in-motion”, Journal of Mechanical Engineering Science.