

Article

# Impact Location and Quantification on an Aluminum Sandwich Panel Using Principal Component Analysis and Linear Approximation with Maximum Entropy

Viviana Meruane \*, Pablo Véliz, Enrique López Droguett and Alejandro Ortiz-Bernardin

Department of Mechanical Engineering, Universidad de Chile, Beauchef 851, Santiago 8370456, Chile; pabloed.veliz@gmail.com (P.V.); elopezdroguett@ing.uchile.cl (E.L.D.); aortizb@ing.uchile.cl (A.O.-B.)

\* Correspondence: vmeruane@ing.uchile.cl; Tel.: +56-2-29784597

Academic Editor: Dawn E. Holmes

Received: 4 January 2017; Accepted: 19 March 2017; Published: 25 March 2017

**Abstract:** To avoid structural failures it is of critical importance to detect, locate and quantify impact damage as soon as it occurs. This can be achieved by impact identification methodologies, which continuously monitor the structure, detecting, locating, and quantifying impacts as they occur. This article presents an improved impact identification algorithm that uses principal component analysis (PCA) to extract features from the monitored signals and an algorithm based on linear approximation with maximum entropy to estimate the impacts. The proposed methodology is validated with two experimental applications, which include an aluminum plate and an aluminum sandwich panel. The results are compared with those of other impact identification algorithms available in literature, demonstrating that the proposed method outperforms these algorithms.

**Keywords:** impact identification; barely visible impact damage; principal component analysis; linear approximation; maximum entropy; sandwich panel

---

## 1. Introduction

The high stiffness and strength at a minimum weight make sandwich structures attractive for applications where weight-saving is critical. Consequently in recent years, the applications of sandwich structures have been rapidly increasing and range from satellites, spacecraft, aircraft, ships, automobiles, rail cars, wind energy systems and bridge construction, among others [1]. Nevertheless, despite previous advantages, these structures can experience damage due to impact loads. To avoid catastrophic failures, it is of critical importance to detect the presence of an impact damage as soon as it occurs. However, this type of damage, known as barely visible impact damage (BVID), is usually internal and there are no visible indications of its presence on the surface.

BVID can be detected by non-destructive testing (NDT) techniques such as X-ray, c-scan or visual inspections. Nevertheless, these techniques are time-consuming, require access to the portion of the structure being inspected and can only be performed when the system is out of service, which can be impractical in some cases. Impact identification methodologies have been proposed as a complement to NDT. These methodologies would continuously monitor the structure, detecting, locating, and quantifying impacts as they occur. For impact damage, its extent is correlated with the impact energy. Therefore, by locating and quantifying impacts over a structure it is feasible to predict possible damage locations, which allows for scheduling of inspections only when they are necessary, and to perform localized searches, saving inspection time.

Impact identification methodologies can be divided into two groups: model-based and data-driven algorithms. Model-based impact identification involves the solution of a nonlinear inverse problem that requires the evaluation of the numerical model several times [2,3], which can

be exceedingly slow for real-time applications. In addition, model-based algorithms rely on the precision of the numerical model, and any error in the numerical model will be interpreted as an impact identification error. Within data-driven algorithms, methodologies based on classification, pattern recognition and regression have been proposed, with artificial neural networks (ANNs) being most frequently used. Worden and Staszewski [4] and Staszewski et al. [5] used two feed-forward multi-layer perceptron (MLP) networks to identify impacts on a composite plate. The first network was trained to detect the impact location, whereas the second network quantifies the impact magnitude. Haywood et al. [6] investigated two approaches to locate impacts in a composite panel with embedded piezoceramic sensors: MLP network and GA-based triangulation. They concluded that both approaches provide a similar degree of accuracy. LeClerc et al. [7] applied a two-step impact detection algorithm to an aircraft component. First, a classification network finds the region of the impact and afterwards another network localizes its position. With this methodology the researchers were able to obtain better results than those of a single neural network. Sharif-Khodaei et al. [8] trained a neural network to detect impact location on a composite-stiffened panel. The network was trained and tested with data obtained from a finite element model of the structure and the numerical model was validated with experimental data.

Although ANNs can process data very quickly, the slow learning speed and the large number of parameters that need to be tuned within the training stage are drawbacks in their application. In the parametric study performed by Sharif-Khodaei et al. [8] it was demonstrated that the performance of a neural network for impact localization strongly depends on the network architecture (number of layers and number of hidden nodes) and network properties (transfer functions and training algorithm). In addition, ANNs have the disadvantage of over-fitting and getting stuck in local minimum. An alternative are support vector machines (SVMs), which exhibit the advantage of global optimization and higher generalization capabilities than ANNs. Least squares support vector machines (LS-SVMs) further simplify the regression to a problem that can be solved from a set of linear equations [9]. Xu [10] implemented an LS-SVM to locate and quantify impacts in an aluminum plate and the results are compared with those of an ANN approach, demonstrating that LS-SVMs reach more accurate results. Fu and Xu [11] proposed a two-layer SVM to predict the location of impacts on an aluminum plate structure. Input data is obtained from a principal component analysis (PCA) of the strain time signal of piezoelectric sensors located over the plate. The results are compared with those of an ANN, concluding that SVM are capable of accomplishing better impact localization accuracy than ANN. To overcome the slow learning speed of ANN, Huang et al. [12] proposed a new learning algorithm called the extreme learning machine (ELM), which is suitable for single-layer feed-forward networks. This algorithm provides good generalization at fast learning speeds, and the only parameter that needs to be tuned is the number of hidden nodes. Xu [13] compared the algorithm performance among a basic ELM, kernel-ELM and LS-SVMs to localize impacts in an aluminum plate. Xu concluded that kernel-ELM is as precise as LS-SVMs, with lower training and evaluation times. Fu et al. [14] implemented a kernel-ELM for impact localization using PCA to extract features. Results in accuracy are similar to those of the SVM, but the kernel-ELM is faster, making it suitable for real-time applications. The time-reversal approach as been presented as a precise alternative for impact localization [15–18]. This a one-class nearest neighbor algorithm, in which the correlation between different impacts is measured by the signals convolution.

Meruane and Ortiz-Bernardin [19] presented a supervised learning algorithm that uses a linear approximation handled by a statistical inference model based on the maximum-entropy principle, denoted linear approximation with maximum entropy (LME). The merits of this approach are threefold: the selection of only one parameter is needed, the regression is determined after solving a convex optimization problem that has a unique solution, and data is processed in a period of time that is comparable to the one of other regression algorithms such as ANNs and SVMs. In addition, LME does not require classical training in which the algorithm is trained once and then the training data is discarded, but performs a linear approximation using

the training data to estimate the impact. The main advantage is that new data can be easily incorporated into the training database with no need for re-training the algorithm as in the case of ANN and SVM. The LME algorithm was originally developed for damage assessment [19,20], and Sanchez et al. [21] implemented it in impact identification demonstrating that, in this case, LME provides a better performance than ANNs and SVMs.

The principle of an impact identification algorithm is to detect, locate and quantify an impact force with the use of a passive system consisting of piezoelectric sensors distributed over the structure. Nonetheless, the amount of strain–time data collected by the sensors is too large to be used directly in a classification or regression algorithm. Therefore, preprocessing of the data is necessary to extract relevant features. In the literature, different features extracted from the signals in the time domain have been studied. Some example are the time of arrival, the time and amplitude of the first peak, and the time and amplitude of the maximum of the signal, among others [21]. Fu et al. [11,14] showed that by using PCA to extract features from the time signals the impact localization results are greatly improved.

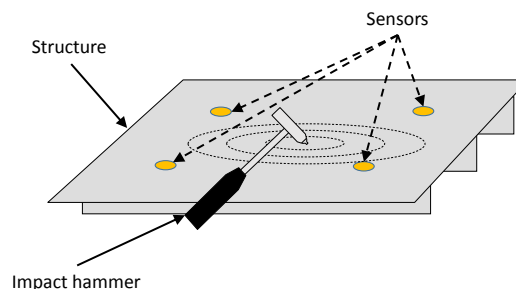
The primary contribution of this research is the development of a novel impact identification algorithm that uses LME in conjunction with PCA. The algorithm implemented by Sanchez et al. [21] uses features extracted from the time signal such as the time of arrival, signal amplitude and information related to the first peak. Therefore, the main differences between the proposed algorithm and the one presented in [21] are with respect to the input features. In addition, our algorithm is evaluated with a stiffened aluminum composite panel that represents a more challenging impact identification problem than the aluminum plate used in references [11,14,21]. The results are compared with those of other impact identification algorithms available in literature [10,11,13–15,21], demonstrating that the proposed method outperforms these algorithms.

The remainder of this work is structured as follows: Section 2 presents the proposed impact identification methodology. Section 3 describes the experimental applications and results. Section 4 compares the results that are delivered by our proposed method with those of other impact identification algorithms available in literature. Finally, conclusions and forthcoming work are presented in Section 5.

## 2. Impact Identification Methodology

### 2.1. Data Acquisition

Figure 1 illustrates a typical experimental setup for impact identification, where piezoelectric sensors bonded to the structure detect surface stress waves generated by the impact. Figure 2 shows an example of a strain–time signal obtained from an impact test and the signal envelope obtained through the Hilbert transform [22].



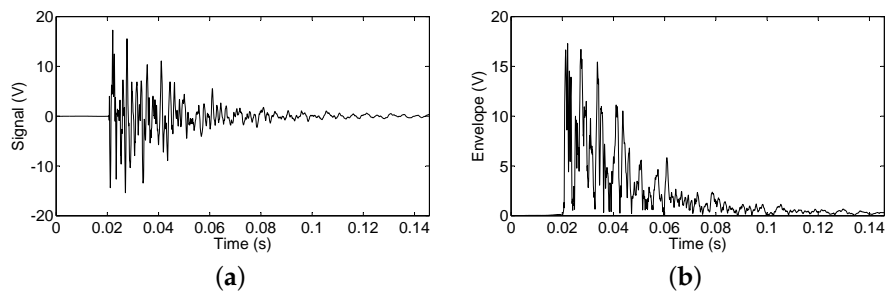
**Figure 1.** Example of an experimental setup for impact identification.

The training data set consists on the time response envelope of the sensors at impacts in different locations in the structure. Let us consider  $n_s$  piezoelectric sensors distributed over the structure, which measure the response at  $q$  training impacts. During an impact the time response of each sensor is recorded in  $p$  data points and the envelope is computed through the Hilbert transform. The response

data corresponds to the voltage provided by the sensors and all sensors are identical. Therefore, the response data is in the same scale units for all the sensors. The information of the  $i$ -th sensor is arranged in the matrix  $\mathbf{Z}_i \in \mathbb{R}^{p \times q}$  as follows,

$$\mathbf{Z}_i = \begin{bmatrix} z_i^1(1) & z_i^2(1) & z_i^3(1) & \dots & z_i^q(1) \\ z_i^1(2) & z_i^2(2) & z_i^3(2) & & z_i^q(2) \\ z_i^1(3) & z_i^2(3) & z_i^3(3) & & z_i^q(3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_i^1(p) & z_i^2(p) & z_i^3(p) & & z_i^q(p) \end{bmatrix}, \quad (1)$$

where  $z_i^j(k)$  is the  $k$ -th response envelope data point of the  $i$ -th sensor at the  $j$ -th impact. Since the number of data points is too large to be used directly in a regression algorithm, preprocessing of the data using principal component analysis (PCA) is performed.



**Figure 2.** Example of a strain-time signal along with its envelope. (a) Strain-time signal; (b) Signal envelope.

## 2.2. Principal Component Analysis

The objective of PCA is to reduce the dimensionality of a data set, while retaining as much as possible the variation present in the data set. This is achieved by transforming it to a new set of uncorrelated variables known as the principal components (PCs).

PCA is defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection lies on the first PC, the second greatest variance on the second PC, and so on.

Let  $\mathbf{Z}_i \in \mathbb{R}^{p \times q}$  be a data matrix with the strain-time response of the  $i$ -th sensor, as described in Equation (1). In general, it is recommended to normalize the data before performing PCA. This is because some of the variables might have a large variance while others a small one, and the variables with a large variance will dominate the first PCs. Nevertheless, in this case the scale of the variables matter because they are related to the location and magnitude of the impacts and it is important to retain that information. Therefore, we have chosen not to normalize the data.

The goal of PCA is to find a matrix  $\mathbf{P}_i \in \mathbb{R}^{n_i \times p}$ , such that a linear mapping from the original dimension  $p$  to a lower dimension  $n_i$  is provided. A new matrix  $\mathbf{S}_i \in \mathbb{R}^{n_i \times q}$ , called score matrix, is obtained from:

$$\mathbf{S}_i = \mathbf{P}_i \mathbf{Z}_i. \quad (2)$$

The  $n_i$  rows of  $\mathbf{P}_i$  are the principal components of  $\mathbf{Z}_i$ .  $\mathbf{P}_i$  can be calculated by extracting the main  $n_i$  eigenvectors of the covariance matrix of  $\mathbf{Z}_i$ ,  $\mathbf{C}_i$ , which is given by:

$$\mathbf{C}_i = \frac{1}{n-1} \mathbf{Z}_i \mathbf{Z}_i^T. \quad (3)$$

The score matrix,  $\mathbf{S}_i$ , contains the features associated with the  $i$ -th sensor. The feature matrix,  $\mathbf{S} \in \mathbb{R}^{n \times q}$ , is obtained by assembling the features from all the sensors as:

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_1 \\ \mathbf{S}_2 \\ \vdots \\ \mathbf{S}_{ns} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^1 & \mathbf{X}^2 & \mathbf{X}^3 & \dots & \mathbf{X}^q \end{bmatrix} \quad (4)$$

where  $ns$  is the number of piezoelectric sensors and  $n$  is the total number of features. The  $j$ -th column of  $\mathbf{S}$ ,  $\mathbf{X}^j$ , corresponds to the PCs associated with the  $j$ -th impact.

### 2.3. Linear Approximation with Maximum Entropy

The observation vector, defined as:  $\mathbf{Y}^j = \{Y_1^j, Y_2^j, Y_3^j\} \in \mathbb{R}^3$ , contains information related to the impact location and magnitude, where  $Y_1^j, Y_2^j$  are the  $x$  and  $y$  coordinates of the force location in  $mm$  and  $Y_3^j$  is the force magnitude in  $N$ . The feature vector  $\mathbf{X}^j = \{X_1^j, X_2^j, \dots, X_n^j\} \in \mathbb{R}^n$ , which is obtained from the  $j$ -th column of  $\mathbf{S}$  in Equation (4), represents the set of PCs associated to the  $j$ -th observation vector  $\mathbf{Y}^j$ . Assuming that a database with a set of  $N$  impacts is constructed, each impact is characterized by an observation vector and a feature vector. Therefore the database is formed by  $N$  pairs of observation and feature vectors as:  $(\mathbf{X}^1, \mathbf{Y}^1), (\mathbf{X}^2, \mathbf{Y}^2), \dots, (\mathbf{X}^N, \mathbf{Y}^N)$ . The central problem in impact identification is: given a certain feature  $\mathbf{X}$ , estimate the corresponding observation  $\mathbf{Y}$ . The nearest neighbor regression estimate of  $\mathbf{Y}$  is given by:

$$\hat{\mathbf{Y}} = \sum_{j=1}^k w_j(\mathbf{X}) \mathbf{Y}^j, \quad (5)$$

where  $\mathbf{Y}^1, \mathbf{Y}^2, \dots, \mathbf{Y}^k$  are the observation vectors associated with the  $k$  closest neighbors to the test vector  $\mathbf{X}$ , and  $w_1(\mathbf{X}), w_2(\mathbf{X}), \dots, w_k(\mathbf{X})$  are weighting functions. The  $k$  nearest neighbor ( $k$ -NN) algorithm weights each neighbor equally, thus  $w_i(\mathbf{X}) = 1/k$ , for  $i = 1$  to  $k$ . A kernel nearest neighbor algorithm bases weightings on the distance from the test vector  $\mathbf{X}$  to each vector in the database [23].

On the other hand, linear approximation takes the  $N$  feature vectors in the database and uses a linear combination of them to represent  $\mathbf{X}$  as [24]:

$$\mathbf{X} = \sum_{j=1}^N w_j(\mathbf{X}) \mathbf{X}^j, \quad \sum_{j=1}^N w_j(\mathbf{X}) = 1, \quad (6)$$

where  $\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^N$  are the feature vectors in the database set. Once the weighting functions are determined, then  $\mathbf{Y}$  is estimated from Equation (5) with  $k = N$ . Typically, Equation (6) is undetermined and its solution can be tackled via an unconstrained optimization technique of the family of least squares. However, these methods produce some negative weights, which lack physical meaning. An alternative that produces positive weights is obtained via the maximum-entropy (max-ent) variational principle [25].

The notion of entropy in information theory was introduced by Shannon as a measure of uncertainty [26]. Later on, Jaynes [25] postulated the maximum-entropy principle as a rationale means for least-biased statistical inference when insufficient information is available. The maximum-entropy principle is suitable to find the least biased probability distribution when there are fewer constraints than unknowns and is posed as follows:

Consider a set of  $N$  discrete events  $\{x_1, \dots, x_N\}$ . The possibility of each event is  $p_i = p(x_i) \in [0, 1]$  with uncertainty  $-\ln p_i$ . The Shannon entropy  $H(\mathbf{p}) = -\sum_{i=1}^N p_i \ln p_i$  is the amount of uncertainty represented by the distribution  $\{p_1, \dots, p_N\}$ . The least biased probability distribution and the one that has the most likelihood to occur is obtained via the solution of the following optimization problem (maximum-entropy principle):

$$\max_{\mathbf{p} \in \mathbb{R}_+^N} \left[ H(\mathbf{p}) = - \sum_{i=1}^N p_i \ln(p_i) \right], \quad (7a)$$

subject to the constraints:

$$\sum_{i=1}^N p_i = 1, \quad \sum_{i=1}^N p_i g_r(x_i) = \langle g_r(x) \rangle, \quad (7b)$$

where  $\mathbb{R}_+^N$  is the non-negative orthant and  $\langle g_r(x) \rangle$  is the known expected value of functions  $g_r(x)$  ( $r = 0, 1, \dots, m$ ). Comparison between Equations (6) and (7b) clearly reveals the analogy between linear approximation and the constraints to which the maximum entropy is subjected. Thus, on replacing the possibility  $p_i$  by the weighting function  $w_i$ , the function  $g_r(x_i)$  by the feature vector  $\mathbf{X}^i$ , and the known expected value  $\langle g_r(x) \rangle$  by the test vector  $\mathbf{X}$ , we have a method to compute the weighting functions that are needed to completely determine the impact estimate  $\hat{\mathbf{Y}}$  in Equation (5).

The optimization problem in Equation (7) assigns probabilities to every  $x_i$  in the set. Now, assume that the probability  $p_i$  has an initial guess  $m_i$  known as a *prior*, which reduces its uncertainty to  $-\ln p_i + \ln m_i = -\ln(p_i/m_i)$ . An alternative problem can be formulated by using this *prior* in Equation (7) [27]:

$$\max_{\mathbf{p} \in \mathbb{R}_+^n} \left[ H(\mathbf{p}) = - \sum_{i=1}^n p_i \ln \left( \frac{p_i}{m_i} \right) \right], \quad (8a)$$

subject to the constraints:

$$\sum_{i=1}^n p_i = 1, \quad \sum_{i=1}^n p_i g_r(x_i) = \langle g_r(x) \rangle. \quad (8b)$$

In Equation (8), the variational principle associated with  $\sum_{i=1}^n p_i \ln \left( \frac{p_i}{m_i} \right)$  is known as *the principle of minimum relative (cross) entropy* [28,29]. Depending upon the *prior* employed, the optimization problem in Equation (8) will favor some  $x_i$  in the set by assigning more probability to them, and eventually, assigning non-zero probability ( $p_i > 0$ ) to a selected number of  $x_i$  ( $i < N$ ) in the set. It can be easily seen that if the *prior* is constant, the Shannon–Jaynes entropy functional in Equation (7) is recovered as a particular case.

The database that is constructed for the impact identification method is usually formed by a large quantity of data elements. Data elements that are far from the impact estimate are not relevant and only introduce noise into the system. Therefore, we need a method to disregard this type of data elements. To this end, the relative entropy approach is suitable and thus we adopt it for our method. We use the prior function to cut off the data elements that are not relevant for the impact estimate. We also replace the probability  $p_i$  and the discrete event  $x_i$  with the weighting function  $w_i$  and the feature vector  $\mathbf{X}^i$  of the linear approximation problem posed in (8), respectively. Finally, the optimization problem that builds our impact identification algorithm reads:

$$\max_{\mathbf{w} \in \mathbb{R}_+^N} \left[ H(\mathbf{w}) = - \sum_{i=1}^N w_i(\mathbf{X}) \ln \left( \frac{w_i(\mathbf{X})}{m_i(\mathbf{X})} \right) \right], \quad (9a)$$

subject to the constraints:

$$\sum_{i=1}^N w_i(\mathbf{X}) \tilde{\mathbf{X}}^i = \mathbf{0}, \quad \sum_{i=1}^N w_i(\mathbf{X}) = 1, \quad (9b)$$

where  $\tilde{\mathbf{X}}^i = \mathbf{X}^i - \mathbf{X}$  has been introduced as a shifted measure for stability purposes. Different prior distributions can be used, typical ones are: Gaussian, cubic spline, quartic spline or constant [30]. Here we tested the four distributions and the best performance was obtained with a smooth Gaussian,

$$m_i(\mathbf{X}) = \exp(-\beta_i \|\tilde{\mathbf{X}}^i\|^2), \quad (10)$$

where  $\beta_i = \gamma/h_i^2$ ;  $\gamma$  is a parameter that controls the support of the Gaussian prior at  $\mathbf{X}^i$ , and therefore its associated weight function; and  $h_i$  is a characteristic  $n$ -dimensional Euclidean distance between neighbors that can be distinct for each  $\mathbf{X}^i$ . In view of the optimization problem posed in Equation (9) for supervised learning, maximizing the entropy chooses the weight solution that commits the least to any one in the database samples [31].

The solution of the max-ent optimization problem is handled by using the procedure of Lagrange multipliers, which yields [27]:

$$w_i(\mathbf{X}) = \frac{Z_i(\mathbf{X}; \lambda^*)}{Z(\mathbf{X}; \lambda^*)}, \quad Z_i(\mathbf{X}; \lambda^*) = m_i(\mathbf{X}) \exp(-\lambda^* \cdot \tilde{\mathbf{X}}^i), \quad (11)$$

where  $Z(\mathbf{X}; \lambda^*) = \sum_j Z_j(\mathbf{X}; \lambda^*)$ ,  $\tilde{\mathbf{X}}^i = [\tilde{X}_1^i \dots \tilde{X}_N^i]^T$ , and  $\lambda^* = [\lambda_1^* \dots \lambda_N^*]^T$  are the converged lagrange multipliers.

In Equation (11), the Lagrange multiplier vector  $\lambda^*$  is the minimizer of the dual optimization problem posed in Equation (12) [27]:

$$\lambda^* = \arg \min_{\lambda \in \mathbb{R}^N} \ln Z(\mathbf{X}; \lambda), \quad (12)$$

which gives rise to the following system of nonlinear equations:

$$\mathbf{f}(\lambda) = \nabla_{\lambda} \ln Z(\lambda) = - \sum_i^N w_i(\mathbf{X}) \tilde{\mathbf{X}}^i = \mathbf{0}, \quad (13)$$

where  $\nabla_{\lambda}$  stands for the gradient with respect to  $\lambda$ . Once the converged  $\lambda^*$  is found, the weight functions are computed from Equation (11) and the impact force is estimated from Equation (5) with  $k = N$ .

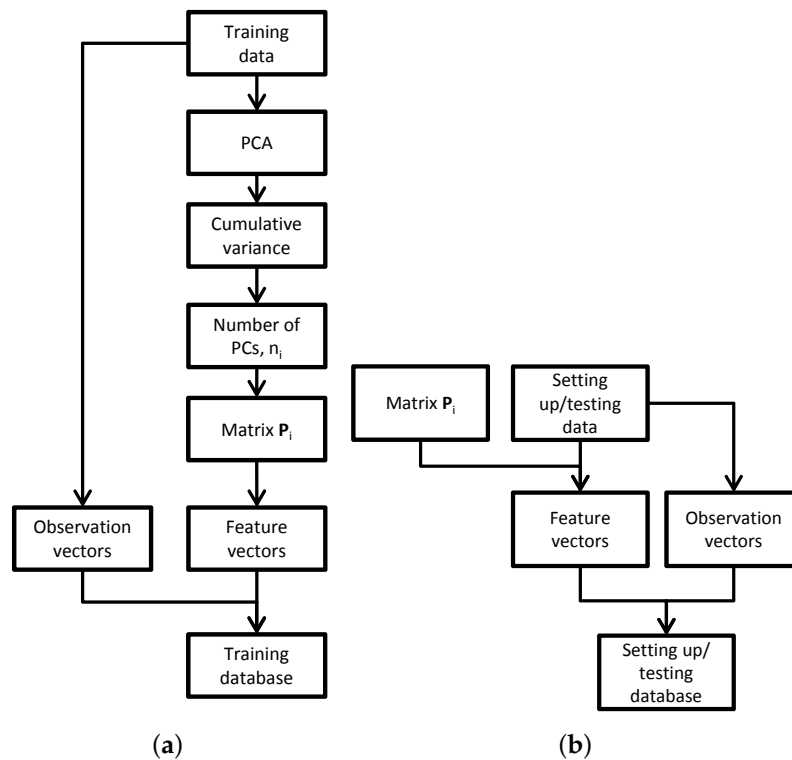
## 2.4. Impact Identification

The impact identification methodology consists of three main parts: building of the databases, selection of parameters and evaluation of the algorithm.

### 2.4.1. Building of the Databases

Three sets of impact data are acquired, one for training, one to set up the parameters of the identification algorithm and one to evaluate the algorithm. In the three cases, each point is impacted once using an instrumented impact hammer. The structure is linear, and as a consequence, the response is proportional to the magnitude of the force. Therefore, the response to impacts of different magnitudes can be determined by simply multiplying the measured response by scaling factors. With this methodology the initial training set is expanded to a new set with impacts of magnitudes between 5 N and 250 N.

Figure 3 presents an scheme of the procedure followed to build the training, setting up and testing databases. In the three cases the information related to the impacts location and magnitude is stored in observation vectors and the time response of the piezoelectric sensors goes into the PCA process described in Section 2.2. First, for the training data, the PCA is performed using  $n_i = p$  and the cumulative percentage variance as a function of the number of PCs is computed. The number of PCs is selected to ensure a cumulative percentage variance of 99.99%. Once  $n_i$  has been selected the matrix  $P_i$  is constructed. Using  $P_i$ , the feature matrices for the training, setting up and testing databases are computed according to Equation (4), from which the feature vectors are extracted (columns of the feature matrices).



**Figure 3.** Construction of the training, setting up and testing databases. (a) Training database; (b) Setting up/testing database. PC: principal components; PCA: principal component analysis.

### 2.4.2. Selection of Parameters

The only parameter that needs to be selected is the number of neighbors that contribute to the solution,  $k$ , which is selected to optimize the performance of the algorithm. To quantify the performance the following error functions are defined:

$$E_x = \frac{1}{n_t} \sum_{j=1}^{n_t} |\hat{Y}_1^j - Y_1^j|, \tag{14}$$

$$E_y = \frac{1}{n_t} \sum_{j=1}^{n_t} |\hat{Y}_2^j - Y_2^j|, \tag{15}$$

$$E_F = \frac{1}{n_t} \sum_{j=1}^{n_t} \frac{|\hat{Y}_3^j - Y_3^j|}{Y_3^j} \times 100, \tag{16}$$

$$E_A = \frac{E_x \times E_y}{A} \times 100 = \frac{\sum_{j=1}^{n_t} |\hat{Y}_1^j - Y_1^j| \sum_{j=1}^{n_t} |\hat{Y}_2^j - Y_2^j|}{n_t^2 A} \times 100, \tag{17}$$



where  $n_t$  is the number of elements in the database;  $A$  is the area of the plate;  $E_x$  and  $E_y$  are the mean errors in the estimation of the force in the  $x$  and  $y$  coordinates;  $E_F$  is the percentage error in the estimation of the force magnitude, and  $E_A$  is the percentage area localization error. The normalized impact identification error is defined as,

$$E_I = \frac{E_A}{E_A^0} \times \frac{E_F}{E_F^0}, \tag{18}$$

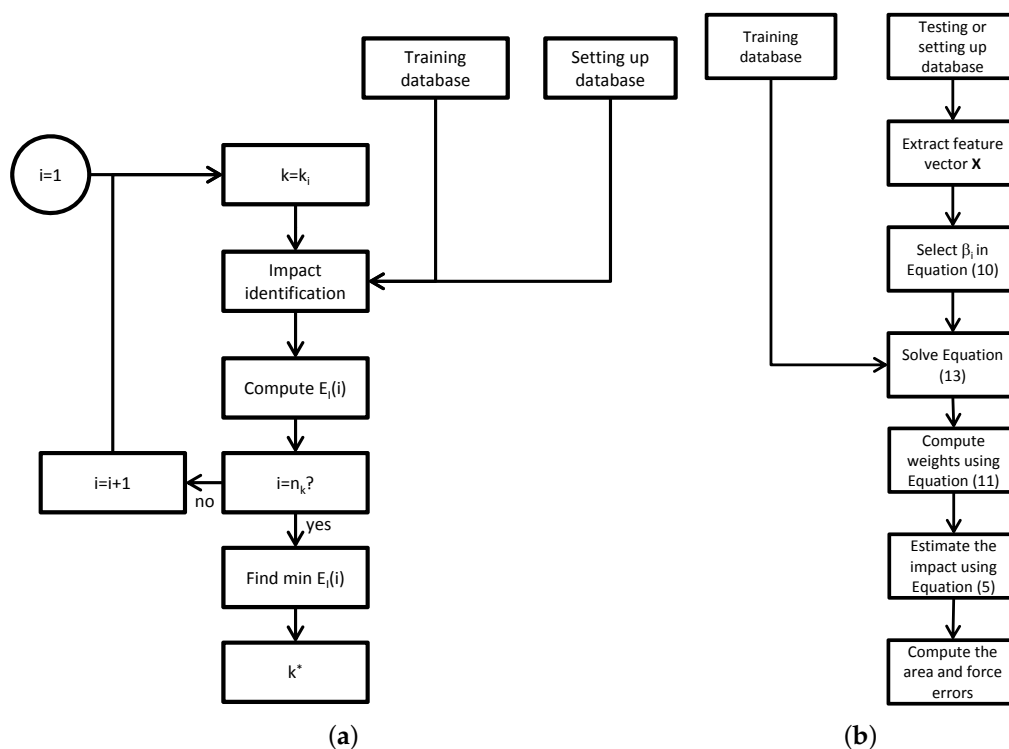
where  $E_A^0$  and  $E_F^0$  are the area and force error for initial value of  $k$ .

Figure 4a illustrates the procedure to select the optimum value of  $k$ . The impact identification algorithm is evaluated for different values of  $k = k_1, k_2, \dots, k_{n_k}$ . The parameter  $k^*$  that provides the lowest normalized impact identification error is selected.

### 2.4.3. Evaluation of the Algorithm

The last step is to evaluate the performance of the algorithm using the testing database. Figure 4b presents the procedure followed to identify the impacts, which consists of the following steps:

1. Extract a feature vector from the testing database.
2. Select the parameter  $\beta_i$  in the Equation (10), so that  $k^*$  neighbors contribute to the solution.
3. Solve the system of nonlinear equations presented in Equation (13).
4. Compute the weight functions using Equation (11).
5. Read the observation vectors in the database and estimate the experimental impact using Equation (5).
6. Compute the area and force errors using Equations (16) and (17).
7. Repeat steps 1 to 6 for all the feature vectors in the testing database.



**Figure 4.** Schemes of the procedure to select the optimum value of  $k$  and to identify an impact. (a) Selection of  $k^*$ ; (b) Impact identification.

### 3. Experimental Applications

#### 3.1. Aluminum Plate

Figure 5 presents the experimental setup, which corresponds to an aluminum plate with dimensions 490 mm × 390 mm × 2.5 mm that is simply supported by four screws. The plate is excited by an instrumented impact hammer and the response is captured by four piezoelectric discs bonded to the surface. Table 1 summarizes the specifications of the experimental equipment.

Data from the four sensors and impact hammer is recorded with a sampling rate of 24 kHz. The hammer is used as trigger and 500 data points before the impact and 3000 data points after the impact are recorded. The signal envelope is computed through the Hilbert transform and no further post processing of the data is performed before PCA.

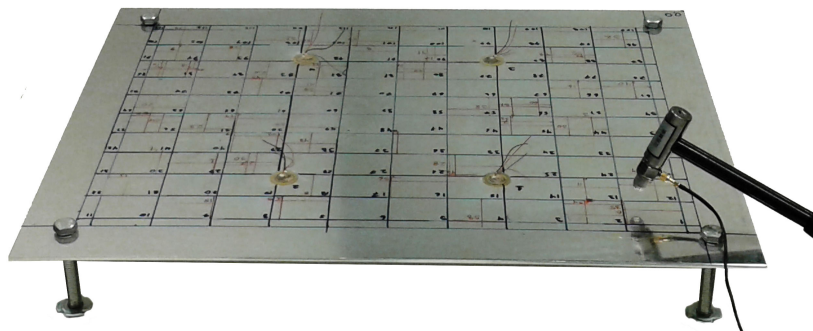


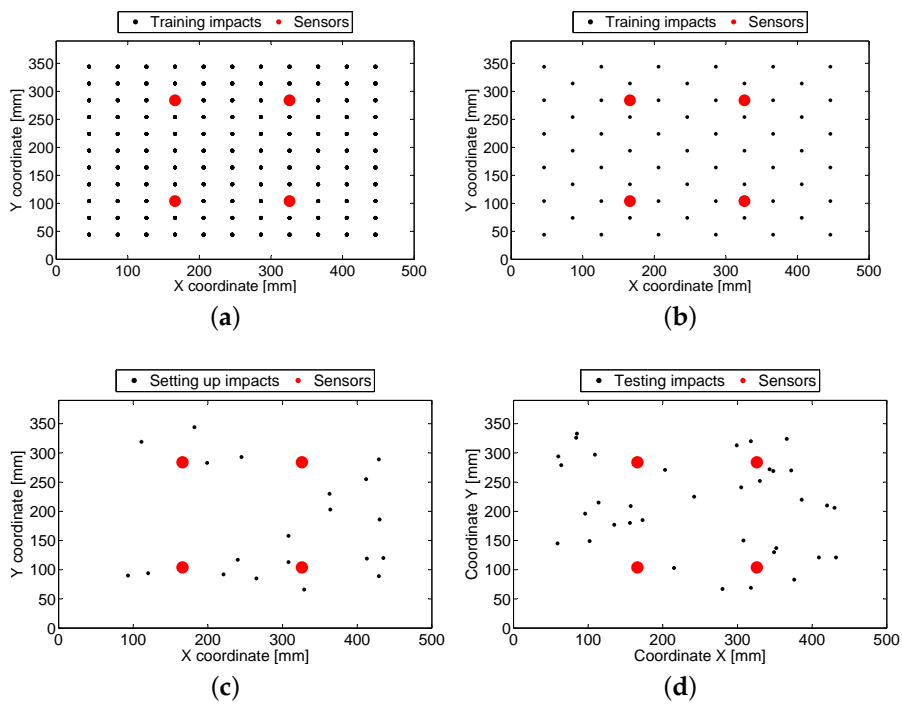
Figure 5. Experimental setup for the aluminum plate.

Table 1. Specifications of the experimental equipment.

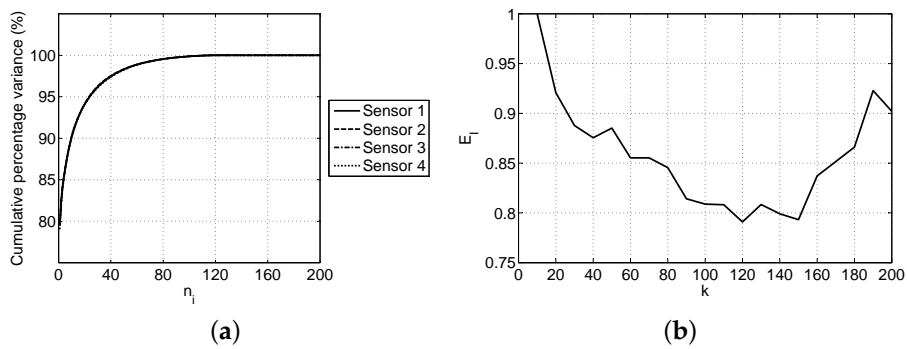
Piezoelectric Discs	
Model	7BB-20-6L0
Resonant frequency	6 kHz
Disc size	20 mm
Thickness	0.42 mm
Impact Hammer	
Model	LC-01A
Sensitivity	4 mV/N
Max. shock force	2 kN
Tip material	Nylon
Force transducer	CL-YD-303
Data Acquisition System	
Model	ECON MI-7016
Resolution	24 bit
Channels	16
Max. sampling rate	96 kHz

Two training sets consisting of a uniform grid of 117 and 61 points are evaluated, as shown in Figure 6a,b, respectively. The setting up and testing sets consist of 20 and 35 random impacts distributed over the plate, as shown in Figure 6c,d, respectively.

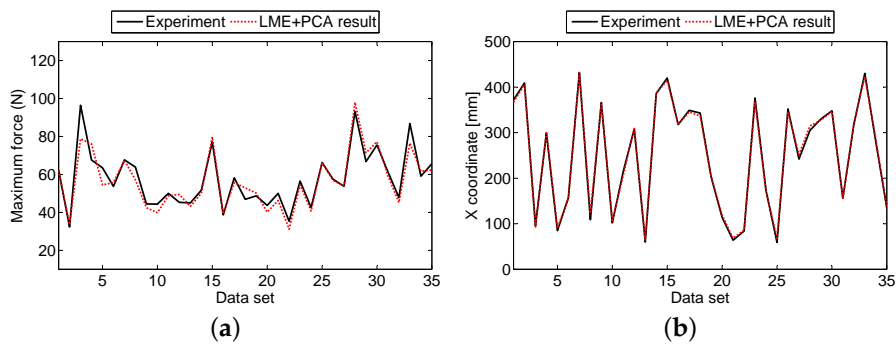
Figure 7a presents the cumulative percentage variance as a function of the number of PCs,  $n_i$ , for the training database 1. With  $n_i = 110$  all sensors have a cumulative percentage variance of 99.99%. Figure 7b shows the normalized impact identification error as a function of  $k$  when  $n_i = 110$ . The minimum error is obtained for  $k = 120$ . Figure 8 presents the impact identification results using these parameters. The area error,  $E_A$ , is 0.009% and the percentage force error,  $E_F$ , is 5.84%.



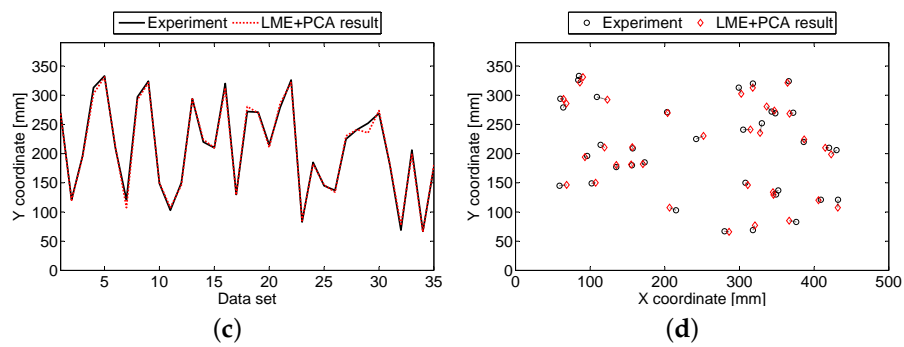
**Figure 6.** Location of the experimental impacts applied to the aluminum plate. (a) Training 1; (b) Training 2; (c) Setting up; (d) Testing.



**Figure 7.** Cumulative percentage variance as a function of  $n_i$  and normalized impact identification error as a function of  $k$  for the first training database. (a) Cumulative percentage variance; (b) Normalized impact identification error.

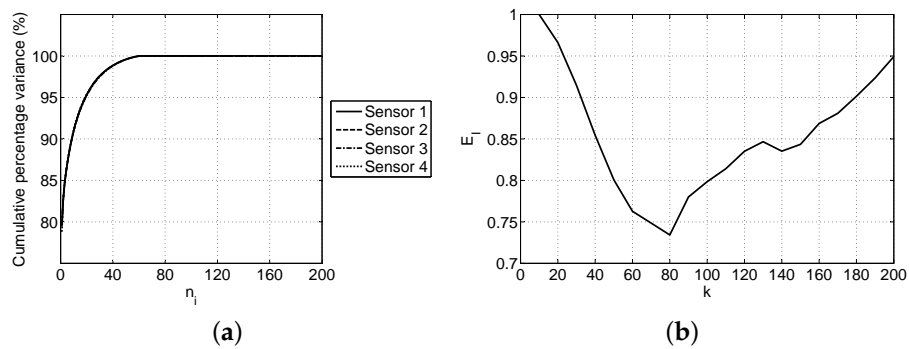


**Figure 8.** Cont.

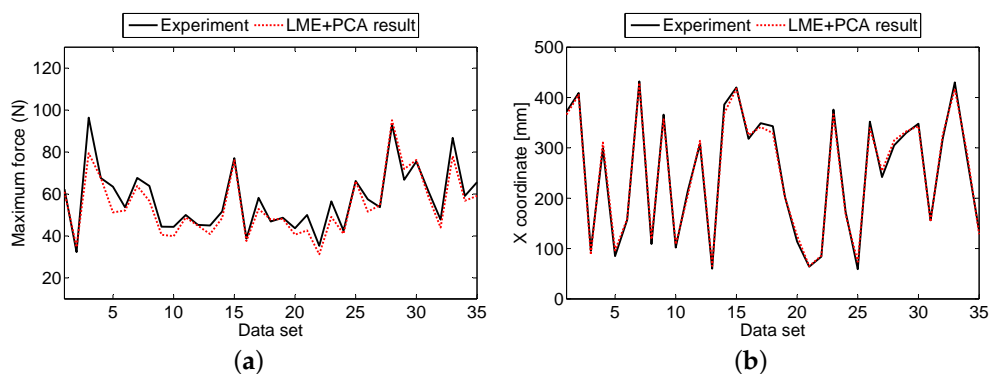


**Figure 8.** Testing results for the aluminum plate using the linear approximation with maximum entropy LME+PCA algorithm and training database 1. (a) Force amplitude; (b) X coordinate; (c) Y coordinate; (d) Localization.

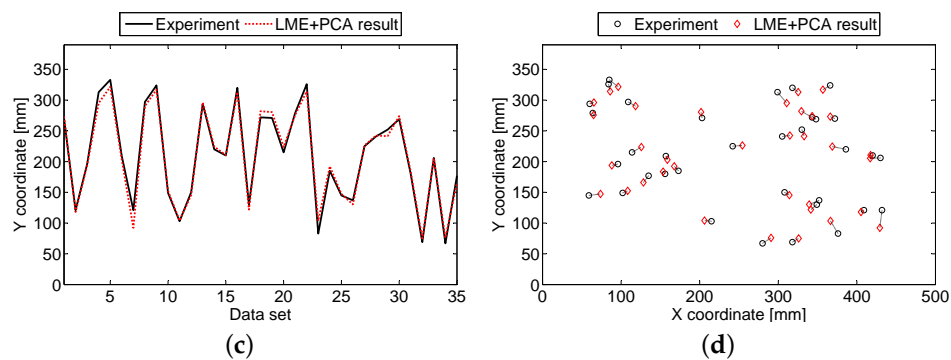
The setting up results for the second training database are presented in Figure 9. The cumulative percentage variance of 99.99% is obtained with  $n_i = 61$  and the minimum normalized impact identification error is obtained with  $k = 80$ . Figure 10 presents the impact identification results using these parameters. The area error,  $E_A$ , is 0.028% and the percentage force error,  $E_F$ , is 6.53%.



**Figure 9.** Cumulative percentage variance as a function of  $n_i$  and normalized impact identification error as a function of  $k$  for the second training database. (a) Cumulative percentage variance; (b) Normalized impact identification error.



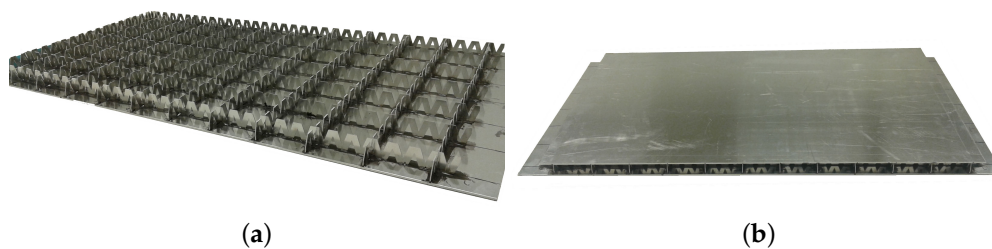
**Figure 10.** Cont.



**Figure 10.** Testing results for the aluminum plate using the LME+PCA algorithm and training database 2. (a) Force amplitude; (b) X coordinate; (c) Y coordinate; (d) Localization.

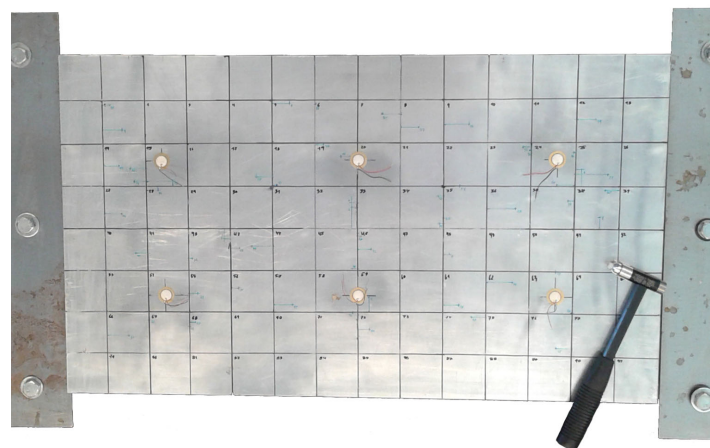
### 3.2. Aluminum Sandwich Panel

The structure consists of a sandwich panel of  $700 \text{ mm} \times 400 \text{ mm} \times 24 \text{ mm}$  made of an aluminum core bonded to two aluminum skins. The core consists of stiffeners with triangular shapes that go through the panel in the two principal directions. The thickness of the skins is 2 mm and the thickness of the stiffeners is 1 mm. The skins are bonded to the stiffeners using an epoxy adhesive cured using a vacuum bagging system. Figure 11a,b show the internal structure and the assembled panel, respectively.



**Figure 11.** Aluminum sandwich panel. (a) Internal structure; (b) Panel.

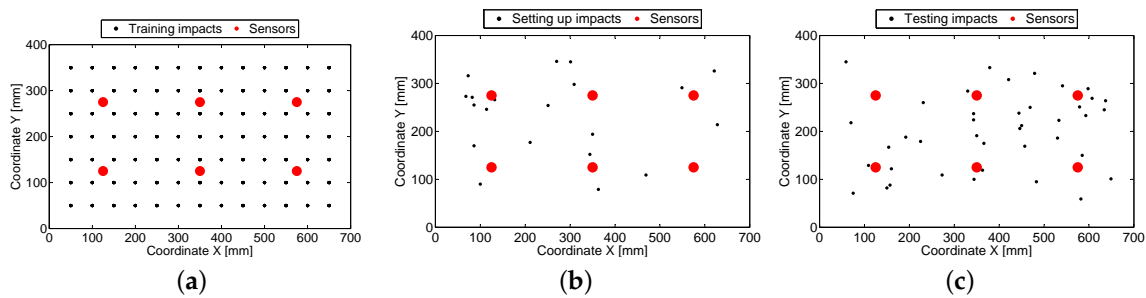
Figure 12 presents the experimental setup. The panel, which is clamped on two edges, is excited by an instrumented impact hammer and the response is captured by six piezoelectric discs bonded to the surface. The specifications of the experimental equipment are presented in Table 1.



**Figure 12.** Experimental setup for the aluminum sandwich panel.

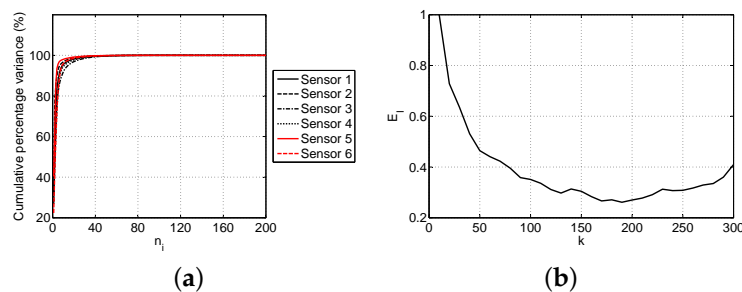
Data from the six sensors and impact hammer is recorded with a sampling rate of 24 kHz. The hammer is used as trigger and 500 data points before the impact and 3000 data points after the impact are recorded. The signal envelope is computed through the Hilbert transform and no further post processing of the data is performed before PCA.

The training set consists of a uniform grid of 91 points, as shown in Figure 13a. The setting up and testing sets consist of 20 and 40 random impacts distributed over the plate, as shown in Figure 13b,c, respectively.

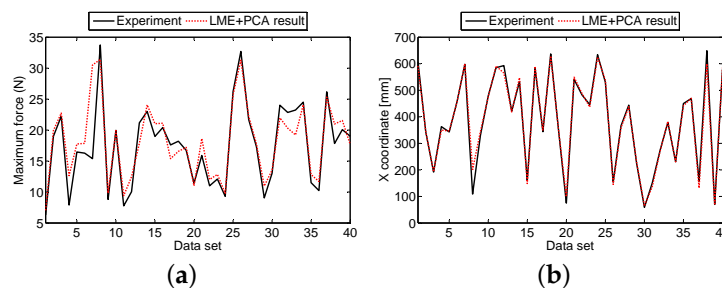


**Figure 13.** Location of the experimental impacts applied to the aluminum sandwich panel. (a) Training; (b) Setting up; (c) Testing.

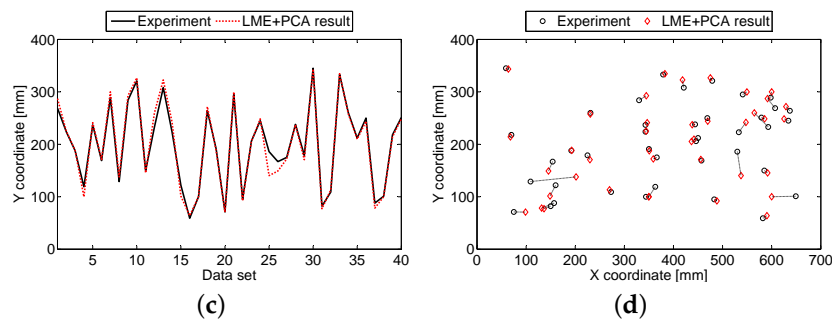
Figure 14a presents the cumulative percentage variance as a function of the number of PCs,  $n_i$ . With  $n_i = 84$  all sensor have a cumulative percentage variance of 99.99%. Figure 14b shows the normalized impact identification error as a function of  $k$  when  $n_i = 84$ . The minimum error is obtained for  $k = 190$ . Figure 15 presents the impact identification results with this combination of parameters. The area error,  $E_A$ , is 0.031 % and the percentage force error,  $E_F$ , is 12.39 %.



**Figure 14.** Cumulative percentage variance as a function of  $n_i$  and normalized impact identification error as a function of  $k$ . (a) Cumulative percentage variance; (b) Normalized impact identification error.



**Figure 15.** Cont.



**Figure 15.** Testing results for the aluminum sandwich panel using the LME + PCA algorithm. (a) Force amplitude; (b) X coordinate; (c) Y coordinate; (d) Localization.

The experimental data and impact identification algorithm for this application case are available for download at [32].

#### 4. Discussion

Table 2 compares the results obtained for the aluminum plate with those of other algorithms available in the literature. These algorithms are evaluated using the same structure and boundary conditions, i.e., an aluminum plate that is simply supported by four screws. The results demonstrate that the proposed methodology has a much better precision than these algorithms.

**Table 2.** Comparison between impact identification algorithms available in the literature and current work. SVM: support vector machine; LS-SVM: least squares SVM; ELM: extreme learning machine.

Reference	Algorithm	Plate Size (mm <sup>2</sup> )	Number of Sensors	Number of Training Impact Points	Area Error (%)	Force Error (%)
Xu [10]	LS-SVM	490 × 390	4	63	1.06	51.2
Fu and Xu [11]	PCA+SVM	490 × 390	4	63	0.13	-
Xu [13]	Kernel-ELM	490 × 390	4	63	0.74	-
Fu et al. [14]	PCA+Kernel-ELM	490 × 390	4	63	0.24	-
Current work	LME+PCA	490 × 390	4	61	0.028	6.53

Furthermore, Table 3 summarizes the testing results for the two experimental cases. In the case of the aluminum plate, the first training database is used. The results are compared with those of the algorithm presented by Sanchez et al. [21]. This algorithm uses LME with features extracted from the time signal such as the time of arrival, signal amplitude and information related to the first peak. Therefore, the only difference are the features used in the LME algorithm. Both algorithms, LME and LME + PCA, have been tested with the same experimental data. The results show that when PCA is used, the precision of the impact identification method has a considerable improvement.

**Table 3.** Performance of the impact identification algorithm.

Experimental Case		Identification Algorithm	
		LME	LME + PCA
Aluminum plate	Area Error (%)	0.12	0.009
	Force Error (%)	7.18	5.84
Sandwich panel	Area Error (%)	0.15	0.031
	Force Error (%)	27.42	12.39

It can also be observed that the results with the aluminum plate are better than with the sandwich panel. This is expected since the stiffeners in the sandwich panel act as obstacles in the propagation of waves, generating reflections that difficult the impact identification procedure.

## 5. Conclusions

This article presented a new impact identification algorithm that uses principal component analysis (PCA) and linear approximation with maximum entropy (LME). The performance of the proposed methodology was validated by considering two experimental cases, which include an aluminum plate and an aluminum sandwich panel. Time varying strain data was measured using piezoceramic sensors bonded to the structures.

To demonstrate the potential of the proposed approach over existing ones, the results were compared with those of other impact identification algorithms available in literature. A much better performance is achieved with our proposed algorithm. The comparison of results using LME with classical time features versus features obtained with PCA shows that with PCA the results are largely improved.

The second experimental case represents a composite panel with stiffeners that can be used for example in aeronautical applications. Therefore, with a low number of sensors it is possible to accurately locate and quantify impacts in realistic structures. Nevertheless, it is necessary to study the performance in structures with more complex geometries.

**Acknowledgments:** The authors acknowledge the partial financial support of the Chilean National Fund for Scientific and Technological Development (Fondecyt) under Grant No. 1160494.

**Author Contributions:** Viviana Meruane and Pablo Véliz designed the experiment, processed and analyzed the experimental data, and tested the impact identification algorithm. Alejandro Ortiz-Bernardin programmed the max-ent linear approximation algorithm and adapted it to the application case. Enrique López Droguett helped with the implementation of the principal component analysis. The article was written by Viviana Meruane, Alejandro Ortiz-Bernardin and Enrique López Droguett.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Vinson, J.R. Sandwich structures: Past, present, and future. In *Sandwich Structures 7: Advancing with Sandwich Structures and Materials*; Springer: New York, NY, USA, 2005.
2. Seydel, R.; Chang, F.K. Impact identification of stiffened composite panels: I. System development. *Smart Mater. Struct.* **2001**, *10*, 354–369.
3. Seydel, R.; Chang, F.K. Impact identification of stiffened composite panels: II. Implementation studies. *Smart Mater. Struct.* **2001**, *10*, 370–379.
4. Worden, K.; Staszewski, W. Impact location and quantification on a composite panel using neural networks and a genetic algorithm. *Strain* **2000**, *36*, 61–68.
5. Staszewski, W.; Worden, K.; Wardle, R.; Tomlinson, G. Fail-Safe sensor distributions for impact detection in composite materials. *Smart Mater. Struct.* **2000**, *9*, 298–303.
6. Haywood, J.; Coverley, P.; Staszewski, W.; Worden, K. An automatic impact monitor for a composite panel employing smart sensor technology. *Smart Mater. Struct.* **2005**, *14*, 265–271.
7. LeClerc, J.; Worden, K.; Staszewski, W.; Haywood, J. Impact detection in an aircraft composite panel—A neural-network approach. *J. Sound Vib.* **2007**, *299*, 672–682.
8. Sharif-Khodaei, Z.; Ghajari, M.; Aliabadi, M. Determination of impact location on composite stiffened panels. *Smart Mater. Struct.* **2012**, *21*, 105026.
9. Gestel, T.V.; Suykens, J.; Baesens, B.; Viaene, S.; Vanthienen, J.; Dedene, G.; Moor, B.D.; Vandewalle, J. Benchmarking least squares support vector machine classifiers. *Mach. Learn.* **2004**, *54*, 5–32.
10. Xu, Q. Impact detection and location for a plate structure using least squares support vector machines. *Struct. Health Monit.* **2014**, *31*, 5–18.
11. Fu, H.; Xu, Q. Locating impact on structural plate using principal component analysis and support vector machines. *Math. Probl. Eng.* **2013**, *2013*, 352149.
12. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: A new learning scheme of feedforward neural networks. In Proceedings of the IEEE International Joint Conference on Neural Networks, Budapest, Hungary, 25–29 July 2004; Volume 2, pp. 985–990.
13. Xu, Q. A Comparison study of extreme learning machine and least squares support vector machine for structural impact localization. *Math. Probl. Eng.* **2014**, *2014*, 906732.



14. Fu, H.; Vong, C.M.; Wong, P.K.; Yang, Z. Fast detection of impact location using kernel extreme learning machine. *Neural Comput. Appl.* **2016**, *27*, 121–130.
15. Ing, R.K.; Quieffin, N.; Catheline, S.; Fink, M. In solid localization of finger impacts using acoustic time-reversal process. *Appl. Phys. Lett.* **2005**, *87*, 204104.
16. Ribay, G.; Catheline, S.; Cloennec, D.; Ing, R.K.; Quieffin, N.; Fink, M. Acoustic impact localization in plates: Properties and stability to temperature variation. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2007**, *54*, 378–385.
17. Park, B.; Sohn, H.; Olson, S.E.; DeSimio, M.P.; Brown, K.S.; Derriso, M.M. Impact localization in complex structures using laser-based time reversal. *Struct. Health Monit.* **2012**, *11*, 577–588.
18. Ciampa, F.; Meo, M. Impact detection in anisotropic materials using a time reversal approach. *Struct. Health Monit.* **2012**, *11*, 43–49.
19. Meruane, V.; Ortiz-Bernardin, A. Structural damage assessment using linear approximation with maximum entropy and transmissibility data. *Mech. Syst. Signal Process.* **2015**, *54*, 210–223.
20. Meruane, V.; Fierro, V.D.; Ortiz-Bernardin, A. A Maximum Entropy Approach to Assess Debonding in Honeycomb Aluminum Plates. *Entropy* **2014**, *16*, 2869–2889.
21. Sanchez, N.; Meruane, V.; Ortiz-Bernardin, A. A novel impact identification algorithm based on a linear approximation with maximum entropy. *Smart Mater. Struct.* **2016**, *25*, 095050.
22. Ulrich, T. *Envelope Calculation From the Hilbert Transform*; Technical Report; Los Alamos National Laboratory: Los Alamos, NM, USA, 2006.
23. Devroye, L.; Györfi, L.; Krzyżak, A.; Lugosi, G. On the strong universal consistency of nearest neighbor regression function estimates. *Ann. Stat.* **1994**, *22*, 1371–1385.
24. Rovatti, R.; Borgatti, M.; Guerrieri, R. A geometric approach to maximum-speed n-dimensional continuous linear interpolation in rectangular grids. *IEEE Trans. Comput.* **1998**, *47*, 894–899.
25. Jaynes, E. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
26. Shannon, C. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
27. Sukumar, N.; Wright, R. Overview and construction of meshfree basis functions: From moving least squares to entropy approximants. *Int. J. Numer. Methods Eng.* **2007**, *70*, 181–205.
28. Kullback, S. *Information Theory and Statistics*; Wiley: New York, NY, USA, 1959.
29. Shore, J.; Johnson, R. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory* **1980**, *26*, 26–36.
30. Arroyo, M.; Ortiz, M. Local maximum-entropy approximation schemes: A seamless bridge between finite elements and meshfree methods. *Int. J. Numer. Methods Eng.* **2006**, *65*, 2167–2202.
31. Gupta, M.R.; Gray, R.M.; Olshen, R.A. Nonparametric supervised learning by linear interpolation with maximum entropy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 766–781.
32. Meruane, V. Source Code for Impact Location and Quantification on an Aluminum Sandwich Panel Using PCA and LME. Available online: [http://viviana.meruane.com/des\\_en.htm](http://viviana.meruane.com/des_en.htm) (accessed on 21 March 2017).



© 2017 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 (<http://creativecommons.org/licenses/by/4.0/>).