Study on Temperature and Synthetic Compensation of Piezo-Resistive Differential Pressure Sensors by Coupled Simulated Annealing and Simplex Optimized Kernel Extreme Learning Machine

Ji Li, Guoqing Hu, Yonghong Zhou, Chong Zou, Wei Peng and Jahangir Alam SM

Article

Abstract: As a high performance-cost ratio solution for differential pressure measurement, piezo-resistive differential pressure sensors are widely used in engineering processes. However, their performance is severely affected by the environmental temperature and the static pressure applied to them. In order to modify the non-linear measuring characteristics of the piezo-resistive differential pressure sensor, compensation actions should synthetically consider these two aspects. Advantages such as nonlinear approximation capability, highly desirable generalization ability and computational efficiency make the kernel extreme learning machine (KELM) a practical approach for this critical task. Since the KELM model is intrinsically sensitive to the regularization parameter and the kernel parameter, a searching scheme combining the coupled simulated annealing (CSA) algorithm and the Nelder-Mead simplex algorithm is adopted to find an optimal KLEM parameter set. A calibration experiment at different working pressure levels was conducted within the temperature range to assess the proposed method. In comparison with other compensation models such as the back-propagation neural network (BP), radius basis neural network (RBF), particle swarm optimization optimized support vector machine (PSO-SVM), particle swarm optimization optimized least squares support vector machine (PSO-LSSVM) and extreme learning machine (ELM), the compensation results show that the presented compensation algorithm exhibits a more satisfactory performance with respect to temperature compensation and synthetic compensation problems.

Keywords: piezo-resistive pressure sensor; temperature compensation; static pressure effect; KELM; CSA; simplex

1. Introduction

Due to the fact pressure usually plays an important role in many industrial processes, a widespread need for pressure sensors has become a reality. One of the most widely used pressure sensors, called piezo-resistive pressure sensor, is made of monocrystalline silicon and fabricated by standard Micro-Electro-Mechanical System (MEMS) craft which conforms to the piezo-resistive principle [1–3]. Since the manufacturing process of the pressure sensor is complicated, some temperature-sensitive defects including inconsistent doping concentration, mismatched thermal expansion coefficients of...
packaging materials and electronics performance may have a dramatic effect on the output if the environmental temperature varies, i.e., they have a temperature effect [4–6]. Besides the temperature effect, the static pressure effect is another key problem that should be taken into consideration when the variable to be measured is differential pressure. Even if the environmental temperature is stable, static pressure can also yield measurement errors in the asymmetric structure of the sensible part of the sensor. Consequently, the input–output characteristics of monocrystalline silicon piezo-resistive pressure sensors becomes non-linear when the environmental temperature or the static pressure change [7].

In order to eliminate the temperature effect in diverse measurement systems, lots of approaches have been proposed, which can be typically categorized into two groups: hardware compensation and software compensation. Hardware compensation modifies the measurement circuits by adding electronical components or develops new structures to achieve more satisfactory performance [8–11]. The complicated debugging process, high cost and relatively limited compensation precision has restricted the spread of hardware compensation applications in engineering practice.

The software compensation approach deems the temperature compensation as a regression problem. The conventional software compensation methods of look-up table, spline interpolation and surface fitting are based on mathematical computation [12–15]. The conventional software compensation approaches seem easy to implement, but when one notices that the compensation quality is proportional to the data scale and the ill-condition dilemma in solving the normalization equation, they become less attractive. On the other hand, their vigorous algorithm robustness and fault tolerance ability has led researchers to focus a great deal of attention on artificial intelligence compensation approaches. The most up-to-date software compensation methods are rooted in artificial intelligence include neural networks [16–19], support vector machines (SVM) and least squares support vector machine (LSSVM) [20–22]. The classic back-propagation neural networks may suffer from the dimensionality curse, local minima, under-fitting or over-fitting, etc. Vapnik [23] developed the SVM, the core of which is the structural risk minimum principle constructed by the empirical risk minimum principle and confidence intervals. Due to the merits of the structural risk minimum principle, the SVM is able to resolve classification and regression problems with relatively small samples. Aiming at alleviating the computational burden of SVM, Suykens [24] proposed the LSSVM, which converts the inequality constraints in traditional SVM into linear equations. Huang et al. [25] presented a fast learning single layer network called extreme learning machine (ELM), wherein the weights and biases of the input layer are randomly assigned within the range of (0, 1) and the weights of the output layer are calculated by a pseudo inverse operation. In combination of the regularization theory, kernel trick and the ELM, the kernel extreme learning machine (KELM) is not only a computationally efficient network, but also has an elegant generalization ability [26].

Although a lot of schemes have been put forward to tackle the issue of temperature compensation, approaches related to synthetic compensation of the differential pressure sensor which consider both temperature and static pressure are rarely reported. In an attempt to compensate the temperature error and static pressure error at the same time, the KELM is adopted in this research. The regularization parameter and the spread parameter of the kernel function govern the KELM’s actual regression precision, hence a parameter selection strategy like the coupled simulated annealing (CSA) hybrid Nelder-Mead simplex searching is employed. Finally, a well-trained synthetic compensation model for differential pressure measurement under different static pressure and temperature conditions is established.

2. Temperature Effect and Static Pressure Effect

The piezo-resistive pressure sensor used in this research is packaged as shown in Figure 1. The inside of the packaged differential pressure sensor is filled with a low thermal expansion rate silicon oil, which serves as the pressure transmission media. The external pressure is applied on the surface of the isolated steel membrane and transferred to the sensor by the silicon oil. The differential pressure \( P_d \) can be calculated by the pressure on the L side \( P_L \) and the pressure on the H side \( P_H \) of the packaged sensor, which is \( P_d = P_H - P_L \).
2.1. Temperature Effect

The configuration of the piezo-resistive pressure sensor is a symmetrical Wheatstone bridge
\((R_1 = R_2 = R_3 = R_4 = R)\), which is illustrated in Figure 2. By analyzing this Wheatstone bridge the
output voltage can be expressed as follows [27]:

\[
U_{\text{output}} = I_f R_0 \pi_0 [1 + (\alpha + \beta + \alpha \beta \Delta T) \Delta T] \left[ \sigma_0 + \frac{\alpha_s - \alpha_g E_0 \Delta T}{1 + \mu} + \theta \right]
\]  
(1)

where \(I_f\) is a constant current supplied by a steady current source, \(R_0\) and \(\pi_0\) are the resistance value
and piezo-resistive coefficient at room temperature. The contents of the first bracket represent the
temperature influence, where \(\alpha\) and \(\beta\) are the temperature coefficients of \(R\) and \(\pi\), respectively, \(\Delta T\)
is the change in the environmental temperature value. Three items in the second bracket stand for
the total force applied to the sensor, where the first item \(\sigma_0\) is the original stress applied on the sensor,
the second item characterizes the additional stress when the ambient temperature varies, \(\alpha_s\) and \(\alpha_g\)
are the thermal expansion coefficients of silicon and glass, respectively, and \(E_0\) is the temperature
coefficient of silicon at the Kelvin temperature, \(\mu\) devotes the Poisson ratio of the monocrystalline
silicon, the third item \(\theta\) is an offset item that illustrates the fabrication inconsistencies between every
bridge arm resistance.

\[
\text{Figure 1. Schematic of the packaged piezo-resistive differential pressure sensor.}
\]

\[
\text{Figure 2. Wheatstone bridge.}
\]

It can be observed from Equation (1) that the actual bridge resistance and piezo-resistive coefficient
directly depend on the environmental temperature. Furthermore, the thermal expansion discrepancy
between the sensor cup and the glass base becomes non-ignorable as the environmental temperature
varies a lot. The relationship clearly unveils the essence of the nonlinearity characteristic between
input pressure and output voltage when the environmental temperature changes.
2.2. Static Pressure Effect

The silicon cup is glued on the glass base, the cross section of which is shown in Figure 3. The H side of the silicon cup is set as the reference side and the pressure acting on this side is called the static pressure. It can be seen that the pressure in the horizontal direction can be offset by the structural symmetry of the silicon cup. Nevertheless, the pressure in the vertical side would cause the deformation of the silicon cup which means a change of the Wheatstone bridge’s arm resistance. Then the output of the sensor is determined according to the Wheatstone bridge state. There are two main aspects that can be taken as the reasons that cause the static pressure effect. On the one hand, there are distinct differences between the mechanical characteristics of monocrystalline silicon and glass. On the other hand, the overall structure formed by the silicon cup and the glass base is not symmetrical. Considering the factors mentioned above, the silicon cup and the glass base may deform quite differently from each other under the same pressure conditions, which will result in a nonlinear output of the sensor. Generally, the static pressure is higher than the differential pressure in actual situations. Moreover, the more the pressure applied on the packaged pressure sensor the more severely the mismatch deformation between the silicon cup and the glass base is. High static pressure may force the sensor working out of the ideal working pressure range. Obviously, to maintain the same high measurement precision as in the condition without static pressure, it is necessary to perform synthetic compensation for occasions of high static pressure.

![Figure 3. The cross section of the differential pressure sensor.](image)

3. Kernel Extreme Learning Machine

As a kind of single layer feedforward network, its structure simplification and computational efficiency endow the ELM with the ability to sufficiently approximate the actual mapping structure embedded in the sample at an extremely fast speed. Suppose an independent and identical distributed sample including N data points as \((x_1, y_1), (x_2, y_2), (x_N, y_N) \in \mathbb{R}^n \times \mathbb{R}\) where \(x_i\) is input, \(y_i\) is output. The details of the approximation function can be expressed as:

\[
\begin{align*}
\sum_{j=1}^{L} \beta_1 h(\omega_1 x_1 + b_1) &= y_1, \\
&\quad \ldots \\
\sum_{j=1}^{L} \beta_j h(\omega_j x_i + b_j) &= y_i, \\
&\quad \ldots \\
\sum_{j=1}^{L} \beta_L h(\omega_L x_N + b_L) &= y_N.
\end{align*}
\]

(2)

where \(L\) is the number of the hidden layer, \(\beta_j\) represents the output layer weight links the \(j\)-th hidden node to the output, \(\omega_j\) and \(b_j\) denote the input layer weights and bias of the \(j\)-th hidden node,
respectively, \( h(\cdot) \) is the activation function for ELM. Furthermore, Equation (2) can be written in matrix form:

\[
H \beta = Y
\]  

(3)

where \( H \) is the hidden layer computation matrix: \( H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_2) \end{bmatrix} \), \( \beta \) is the output layer weight vector and \( Y \) is the output matrix. Given the randomly generated input layer weights and the specified activation function, the output layer weight vector can be calculated in terms of the norm least square method. Then, the output layer weight vector can be represented as follows:

\[
\beta = H^+ Y
\]

(4)

where \( H^+ \) is the Moore-Penrose generalized inverse of \( H \).

Huang et al. proposed the unified ELM in pursuit of a more unified framework for classification and regression. According to the regularization theory, a positive item \( 1/C \) (where \( C \) is the penalty parameter) is incorporated into the calculation of the output weight vector and Equation (4) turns into:

\[
\beta = H^T (HH^T + \frac{1}{C})^{-1} Y
\]

(5)

Meanwhile, the regression results with respect of inputs can be defined as:

\[
f(x) = HH^T \left( HH^T + \frac{1}{C} \right)^{-1} Y
\]

(6)

\[
HH^T = \begin{bmatrix}
h(x_1)h(x_1) & \cdots & h(x_1)h(x_j) & \cdots & h(x_1)h(x_N) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
h(x_i)h(x_1) & \cdots & h(x_i)h(x_j) & \cdots & h(x_i)h(x_N) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
h(x_N)h(x_1) & \cdots & h(x_N)h(x_j) & \cdots & h(x_N)h(x_N)
\end{bmatrix}
\]

(7)

In light of the Cover theorem, inputs in a nonlinear separable space become easier to separate by mapping them into a higher dimensional space. The kernel trick is always the first choice for the implementation of this work, the KELM thus uses it to replace the elements in Equation (7). So Equation (7) can be rewritten as:

\[
\Omega_{KELM} = HH^T = \begin{bmatrix}
K(x_1,x_1) & \cdots & K(x_1,x_j) & \cdots & K(x_1,x_N) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
K(x_i,x_1) & \cdots & K(x_i,x_j) & \cdots & K(x_i,x_N) \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
K(x_N,x_1) & \cdots & K(x_N,x_j) & \cdots & K(x_N,x_N)
\end{bmatrix}
\]

(8)

Finally, the KELM regression model can be denoted as:

\[
f(x) = \begin{bmatrix}
K(x,x_1) \\
\vdots \\
K(x,x_j) \\
\vdots \\
K(x,x_N)
\end{bmatrix}^T \left( \Omega_{KELM} + \frac{1}{C} \right)^{-1} Y
\]

(9)
Sensors 2017, 17, 894

Since the kernel function is one of the critical factors that contribute to approximation capability, a reasonable choice of kernel function is necessary. Actually, kernel that meets the Mercer kernel condition, namely admissible kernel, can be set as a kernel function. A typical and extensively used kernel function is the Radial Basis Function (RBF) kernel, which is defined as [28]:

\[
K_{\text{RBF}} = \exp\left(-\frac{\| x - x_i \|^2}{2\sigma^2}\right)
\]

(10)

where the spread \(\sigma\) is the only dependent kernel parameter to be selected. However, the proposed scheme is also feasible with other kernels in a same manner.

4. Coupled Simulated Annealing and Simplex Search

4.1. Coupled Simulated Annealing (CSA)

Noting that the setting of parameters in KELM is vital to the final model performance, the coupled simulated annealing (CSA) optimization method proposed by Xavier de Souza et al. [29] is employed to perform the parameter search. Like a previously published algorithm, simulated annealing [30] simulates the metal annealing process. According to statistical mechanics, the distribution of atom energy in the \(i\)-th state at a specified temperature level conforms to the Boltzmann formulation:

\[
P(E = E_i) = \frac{1}{Z} \exp\left(-\frac{E_i}{k_B}\right)
\]

(11)

where \(E_i\) is the energy in the \(i\)-th state, \(k_B\) is the Boltzmann constant, \(Z\) is the partition function. The simulated annealing is a kind of heuristic algorithm that iterates in the decreasing direction of object function. The Metropolis principle also allows the iteration process to accept the increasing object function value with a relative small probability which contributes to the dynamics to jump out of the local minimum. The coupled simulated annealing is developed from the classical simulated annealing but the parallel mechanism within it is distinct from the latter. The coupled simulated annealing introduced a multi-start initialization into the probing process over the solution space to supply more a priori information. The coupled item is created to help decide if the less favorable solution is accepted and to simplify the computational complexity. The coupled item and acceptance probability take the following form:

\[
A_\Theta(\gamma; x_i \to y_i) = \frac{\exp\left(\frac{E(x_i)}{T^\text{acc}_k}\right)}{\gamma}
\]

(12)

\[
\gamma = \sum_{x_i \in \Theta} \exp\left(\frac{E(x_i)}{T^\text{acc}_k}\right)
\]

(13)

where \(x_i\) and \(y_i\) are states belonging to the possible state space, \(E(x_i)\) denotes the energy in \(x_i\), \(T^\text{acc}_k\) is the acceptance temperature at the \(k\)-th iteration, \(\Theta\) represents the current state set, \(A_\Theta(\gamma; x_i \to y_i)\) illustrates the acceptance probability from state \(x_i\) to state \(y_i\) with the coupled item \(\gamma\). The coupled simulated annealing algorithm can be depicted as follows:

1. Initialization: M random solutions are assigned to \(\Theta\). Evaluate the energy \(E(x_i)\) and coupled term \(\gamma\). Set \(T_k\) as \(T_0\), \(T^\text{acc}_k\) as \(T^\text{acc}_0\) and the iteration index \(k\) equals to zero. The variance of the acceptance probability is calculated according to \(\sigma^2_\Theta = 0.99(m - 1/m^2)\), where \(m\) is the number of all states included in \(\Theta\). The rate controls the temperature variation marked as \(\alpha\) is set to 0.05.

2. A new state \(y_i\) is generated that corresponds to the current state \(x_i\) in the state space by \(y_i = x_i + \epsilon_i\) where \(\epsilon_i\) is independently and randomly sampled from a normal distribution at temperature \(T_k\) as \(g(\epsilon_i, T_k)\). Evaluate all the energy values \(E(y_i), \forall i = 1, \cdots, m\).
3. The new state is accepted if \( E(y_i) \leq E(x_i) \) or a random number \( \text{rand} \) is generated by a uniform distribution in [0, 1] and compared with the acceptance probability \( A_Q(\gamma; x_i \rightarrow y_i) \) obtained by Equation (12) when it satisfies the condition \( A_Q > \text{rand} \). Otherwise, the old state remains. Assess \( \gamma \) and return to step 2 to achieve thermal equilibrium condition.

4. Adjust the acceptance temperature \( T^{\text{acc}}_k \) by the rules: if \( \sigma^2 < \sigma^2_D \), \( T^{\text{acc}}_k = (1 - \alpha) \times T^{\text{acc}}_{k-1} \), if \( \sigma^2 > \sigma^2_D \), \( T^{\text{acc}}_k = (1 + \alpha) \times T^{\text{acc}}_{k-1} \).

5. Increment the annealing time \( k \) and decrease the temperature according to the annealing scheme: \( T_k = T_0 / k \).

6. Terminate the iteration if the current energy value meets the stopping criterion, otherwise, go back to step 2.

4.2. Simplex Search

The simplex search was developed by Nelder and Mead [31] as a simplex is a geometric figure that has \( n + 1 \) vertices. The expansion operation and the contraction operation along the line of reflection allow the simplex search to considerably reduce the computation time. Assuming the initial simplex \( X = \{x_1, x_2, \cdots, x_{n+1}\} \) so that the output is assigned as \( X \) because it is our new simplex. It should be noted that \( X \) is an \( n \times (n + 1) \) matrix, where each column represents a simplex vertex. The simplex searching process is described as follows:

1. Build \( n + 1 \) vertices of \( X \), evaluate and sort their function values.

2. Compute the reflection point \( x_c \) by \( x_c = \bar{x} + \alpha (\bar{x} - x_{n+1}) \), where \( \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \) is the centroid of the \( n \) points except for the worst point as \( x_{n+1} \). If \( f(x_1) \leq f(x_r) \leq f(x_{n+1}) \), accept the reflected point \( x_r \).

3. If \( f(x_r) < f(x_1) \), perform the expansion operation as \( x_c = \bar{x} + \beta (x_r - \bar{x}) \). If \( f(x_c) \leq f(x_r) \), accept the expanded point \( x_c \); otherwise accept \( x_r \).

4. If \( f(x_r) \geq f(x_{n+1}) \), a contraction should be done utilizing \( \bar{x} \) with the better point of \( x_{n+1} \) and \( x_r \):
   
   (a) If \( f(x_r) < f(x_{n+1}) \), then outside contraction: \( x_c = \bar{x} + \gamma (x_r - \bar{x}) \). If \( f(x_c) < f(x_r) \), accept \( x_c \); otherwise go to step 5.
   
   (b) If \( f(x_r) \geq f(x_{n+1}) \), then inside contraction: \( x_c = \bar{x} - \gamma (x_r - x_{n+1}) \). If \( f(x_c) \leq f(x_{n+1}) \), accept \( x_c \); otherwise go to step 5.

5. Calculate \( f \) at the \( n \) points \( v_i = x_1 + \delta (x_i - x_1), i = 2, \cdots, n + 1 \). The vertices at the next iteration are made up of \( X = \{x_1, v_2, \cdots, v_{n+1}\} \).

The flowchart of the temperature and synthetic compensation algorithm is illustrated in Figure 4. Taking all the merits of CSA and simplex search into account it is reasonable to model the non-linear relationship between the input and output of the piezo-resistive differential pressure sensor as follows:

1. Divide the normalized sample into training set and testing set according to the engineering requirements.

2. Initialize the KELM with a random parameter set in an acceptable range.

3. Initialize the parameters in CSA such as the number of states at a certain temperature \( M \), the coupled term value \( \gamma \), the starting temperature \( T_0 \), the starting acceptance temperature \( T^{\text{acc}}_0 \) and the temperature step regulating rate \( \alpha \).

4. Evaluate the function iteratively until the maximum iteration number is reached or the fitness is less than the limit \( \varepsilon_1 \). A suboptimal hyper-parameter set \( (C_1, \sigma_1) \) is found through this step.

5. The simplex search is performed with the start solution as \( (C_1, \sigma_1) \) until the maximum iteration number is reached or the difference between fitness in two successions is smaller than the limit \( \varepsilon_2 \). Finally, a more satisfactory parameter set as \( (C_2, \sigma_2) \) can be obtained.
5. Experiments and Result Analysis

5.1. Calibration Experiment Setup

A calibration experiment was performed on a 0.065% grade pressure sensor without compensation [32] to provide data for modeling the KELM approach. The measurement range of the pressure sensor used in this calibration is from $-1000$ kPa to $1000$ kPa, and the working temperature range is from $-20$ °C to $70$ °C. The environmental temperature and the differential pressure output, which were marked as $T_{AD}$ and $U_{AD}$, are converted by an analog-digital (A/D) converter. A two channel pressure controller exerts the H side pressure ($P_H$) and the L side pressure ($P_L$).

The static pressure ($P_H$) which is denoted by $P_s$ is sampled at 14 levels ($0$, $1000$, $2000$, $2500$, $3000$, $4000$, $4500$, $5000$, $6000$, $6500$, $7000$, $8000$, $8500$ and $9000$ kPa). The differential pressure denoted as DP is specified in a step of $125$ kPa from $-1000$ kPa to $1000$ kPa at each static pressure level. Data calibration is performed for five temperature preservation processes at temperatures of $-20$, $0$, $20$, $50$ and $70$ °C. Every temperature preservation process takes about $3$ h. The input variables used to construct the KELM include the differential pressure output $U_{AD}$, the static pressure $SP$ and the environmental temperature $T_{AD}$, while the output is the differential pressure $DP$. A total of $17 \times 14 \times 5 = 1190$ input-output pairs were collected when the experiments finished. The experimental setup is shown in Figure 5.

The calibration data results without static pressure ($SP = 0$ kPa) are illustrated as Figure 6a. It can be seen in Figure 6a that the sensor’s output performance is affected by the environmental temperature notably without any static pressure participation. To be more specific, relative errors ($Err_{temp}$) with respect to the linear fit characteristic according to data at $20$ °C are depicted in Figure 6b, which shows the maximum drift value reaches $1.75\%$. On the other hand, the maximum measurement errors caused by the static pressure effect are $0.31\%$, $0.11\%$, $0.16\%$, $0.2\%$ and $0.13\%$ corresponding to temperature settings of $-20$, $0$, $20$, $50$ and $70$ °C, respectively. As demonstrated in Figure 7, the measurement errors increase gradually with the static pressure and also fluctuate randomly at different...
static pressure levels. As a matter of fact, both the silicon cup and the glass base experience distortion to a different degree when the force status and temperature conditions change, which might be the main factor contributing to this phenomenon. To meet the high engineering requirements, a synthetic compensation strategy should be taken into account to address the remarkable temperature and static pressure effects.

**Figure 5.** Setup for the calibration experiments.

(a) Calibration results without static pressure

(b) Relative errors at different temperatures

**Figure 6.** Pressure sensor’s output and relative error at different temperatures.
5.2. Logarithmic Transformation of Dependent Parameters and Normalization

To speed up and facilitate the search process in the parameter space, the penalty parameter C and the kernel parameter $\sigma$ use $\ln$ to accomplish a logarithmic transformation. Additionally, features in the calibration experiment have mismatched levels of magnitude, which would result in an undesirable solution. Thus, the experiment sample should be normalized into the $[-1, 1]$ space according to the form:

$$\hat{x}_i = \frac{x_i - (x_i(\text{max}) - x_i(\text{min}))}{x_i(\text{max}) - x_i(\text{min})}/2,$$

where $x_i$ is the i-th variable of a data point, and $x_i(\text{max})$ and $x_i(\text{min})$ are the upper limit and lower limit of i-th variable, respectively. After normalizing the experimental data, a training set and a testing set should be partitioned before modeling.

5.3. Compensation Result Analysis

To evaluate the compensation effectiveness of the CSA and simplex search optimized KELM (CSA-simplex-KELM) model, several other methods, including BP, RBF, particle swarm optimization optimized SVM (PSO-SVM), particle swarm optimization optimized LSSVM (PSO-LSSVM) and ELM were also investigated. The performance measure of all the compensation algorithms is the absolute relative error, which takes the form of:

$$Err_i = \left| \frac{DP_i' - DP_i}{DP_{FS}} \right| \times 100\%,$$

where $i$ is the index of differential pressure, $DP_i'$ is the compensated differential pressure, $DP_i$ is the reference differential pressure, and $DP_{FS}$ is the sensor full scale.

Because PSO and CSA are heuristic algorithms, the parameter set $(C, \sigma)$ is defined as an individual solution in their search space. To achieve a balance between the regression accuracy and algorithmic stability, the mean squared error (MSE) of $Err$ is taken as the algorithm fitness function. The parameters of each method are tabulated in Table 1. Considering the importance of cross validation for the determination of the model and the computation efficiency, generalized cross validation is adopted to prevent overfitting [33]. The whole coding is implemented with the help of LS-SVMLab toolbox and libsvm toolbox on the MATLAB (2015b) platform (The Mathworks, Inc., Natick, MA, USA) [34,35].
Table 1. Parameters setting of compensation models.

| Parameters                | PSO-SVM | PSO-LSSVM | CSA-Simplex-KELM |
|---------------------------|---------|-----------|------------------|
| swarm size/state level    | 30      | 30        | 6                |
| iteration number/annealing time | 30    | 30        | 30               |
| maximum weight            | 0.9     | 0.9       |                  |
| minimum weight            | 0.4     | 0.4       |                  |
| social factor             | [1, 3]  | 2         |                  |
| cognitive factor          | [1, 3]  | 2         |                  |
| thermal equilibrium steps |         | 5         |                  |
| initial/acceptance temp   |         | 1         |                  |
| regulation rate           |         | 0.1       |                  |
| Penalty parameter (C)     | [1, 1 × 10^7] | [1, 1 × 10^7] | [1, 1 × 10^7] |
| Kernel parameter (σ)      | [1 × 10^{-3}, 10] | [1 × 10^{-3}, 10] | [1 × 10^{-3}, 10] |
| maximum interval tolerance (ε) | [1 × 10^{-6}, 1] |                  |                  |

The differential pressure sensor can work in two ways: with or without static pressure. In the first case the static pressure effect is neglected, which simplifies the task to a temperature compensation problem. From a more practical viewpoint, the second case is more complicated as the temperature and static pressure are coupled. These two cases are discussed in the following subsections.

5.3.1. Temperature Compensation

A training set consists of nine differential pressure levels is uniformly selected from $-1000$ kPa to $1000$ kPa in the working temperature range, which means 45 data pairs in all. The rest of the data without static pressure are put into the testing set. The number of hidden layer nodes except for SVM and LSSVM are determined as in Table 2.

Table 2. Model configuration for temperature compensation.

| Temperature Compensation Methods | Hidden Layer Node Number and Spread Parameter |
|---------------------------------|----------------------------------------------|
| BP                              | 8                                            |
| RBF                             | 37; spread:5.7                               |
| ELM                             | 36                                           |
| CSA-simplex-KELM                | 10                                           |

Four performance indices—minimum (min), maximum (max), mean (mean), variance (var) of the averaged training set errors (Err)—are taken to evaluate the compensation performance of every model, and the details are summarized in Tables 3 and 4. Some conclusions can be directly inferred from the observation of the training set errors in Table 3 and testing set errors in Table 4. The BP neural network does not perform as well as other models in both the training set and testing set, even with the simplest structure, which means it falls into a local optimal solution. The RBF neural network is among the best when merely considering the testing set results, but it may lose some features during the training period, which makes the maximum error comes to $3.8180 \times 10^{-4}$.

Table 3. Temperature compensation results of the training set.

| Temperature Compensation Methods | Err (min)   | Err (max)   | Err (mean)  | Err (variance) |
|---------------------------------|-------------|-------------|-------------|----------------|
| BP                              | $6.1336 \times 10^{-6}$ | $4.4261 \times 10^{-4}$ | $1.1508 \times 10^{-4}$ | $1.3032 \times 10^{-8}$ |
| RBF                             | $3.9938 \times 10^{-6}$ | $3.8180 \times 10^{-4}$ | $8.9686 \times 10^{-5}$ | $5.3639 \times 10^{-9}$ |
| PSO-SVM                         | $6.2186 \times 10^{-6}$ | $3.0263 \times 10^{-4}$ | $1.1508 \times 10^{-4}$ | $5.8987 \times 10^{-9}$ |
| PSO-LSSVM                       | $9.9424 \times 10^{-7}$ | $2.1566 \times 10^{-4}$ | $3.4931 \times 10^{-5}$ | $1.7494 \times 10^{-9}$ |
| ELM                             | $1.0264 \times 10^{-7}$ | $1.2989 \times 10^{-4}$ | $2.0954 \times 10^{-5}$ | $8.1310 \times 10^{-10}$ |
| CSA-simplex-KELM                | $1.5497 \times 10^{-6}$ | $2.3419 \times 10^{-4}$ | $4.5105 \times 10^{-5}$ | $2.0150 \times 10^{-9}$ |
Because the PSO-LSSVM model makes use of all the training data to learn the nature of the mapping relationship, it is not only superior to PSO-LSSVM in all indices, but also gives a rather similar result compared with the ELM. Obviously, the models within the ELM framework can provide a more desirable improvement of measurement accuracy and stability than others.

As seen in the compensation results of all the models, shown in Figure 8, the maximum compensation error occurs at almost 70 °C which explains why the temperature effect is vital to the pressure measurement and difficult to eliminate completely. There is no tendency in any of the compensation results of all the models, which confirms the validity of these models. The PSO-LSSVM, the ELM and the KELM share almost the same global compensation performance, while the proposed KLEM model requires the least number of hidden layer nodes. From the engineering viewpoint, hence, the KELM presents the most satisfactory comprehensive performance in controlling compensation error and convenience of implementation.

Table 4. Temperature compensation results of the testing set.

| Temperature Compensation Methods | Err (min)       | Err (max)       | Err (mean)      | Err (variance)  |
|----------------------------------|-----------------|-----------------|-----------------|-----------------|
| BP                               | $7.3618 \times 10^{-8}$ | $5.0303 \times 10^{-4}$ | $1.4749 \times 10^{-4}$ | $1.6054 \times 10^{-8}$ |
| RBF                              | $2.3279 \times 10^{-8}$ | $2.0774 \times 10^{-4}$ | $7.0203 \times 10^{-5}$ | $3.5305 \times 10^{-9}$ |
| PSO-SVM                          | $3.1682 \times 10^{-6}$ | $2.9965 \times 10^{-4}$ | $1.1040 \times 10^{-4}$ | $5.0093 \times 10^{-9}$ |
| PSO-LSSVM                        | $3.4454 \times 10^{-7}$ | $2.8042 \times 10^{-4}$ | $3.6780 \times 10^{-5}$ | $2.1839 \times 10^{-9}$ |
| ELM                              | $8.0991 \times 10^{-7}$ | $2.5388 \times 10^{-4}$ | $3.2806 \times 10^{-5}$ | $2.2638 \times 10^{-9}$ |
| CSA-simplex-KELM                 | $3.5241 \times 10^{-6}$ | $2.4787 \times 10^{-4}$ | $5.2075 \times 10^{-5}$ | $1.8499 \times 10^{-9}$ |

Figure 8. Temperature compensation results obtained by different models. (a) BP temperature compensation results; (b) RBF temperature compensation results; (c) PSO-SVM temperature compensation results; (d) PSO-LSSVM temperature compensation results; (e) ELM temperature compensation results; (f) KELM temperature compensation results.

5.3.2. Synthetic Compensation

If we note that sensors always work under the influence of the static pressure in most practical environments, a synthetic compensation strategy is critical. Within the working temperature range (−20 °C–70 °C), nine differential pressure levels were uniformly selected from −1000 to 1000 kPa and seven static pressure levels (1000, 2500, 4000, 5000, 6500, 8000 and 9000 kPa) make up the training set.
for the modeling process, which means we have 315 data points. The rest of the calibration data forms the testing set. Except for the PSO-SVM and the PSO-LSSVM, the other models’ performance is related to their topology structure, so the number of hidden layer nodes except for SVM and LSSVM are set as in Table 5.

Table 5. Model configuration for synthetic compensation.

| Temperature Compensation Methods | Hidden Layer Node Number and Spread Parameter |
|----------------------------------|----------------------------------------------|
| BP                               | 8                                            |
| RBF                              | 176; spread:3.7                              |
| ELM                              | 4                                             |
| CSA-simplex-KELM                 | 5                                            |

Like in the temperature compensation, the training set and testing set compensation results are summarized in Tables 6 and 7, respectively. In comparison with the temperature compensation data, the overall quality of the synthetic compensation is reduced to a certain extent. This change proves that the static effect is really vital to the synthetic performance of the sensor. By analyzing the results in Tables 6 and 7, several conclusions can be reached: firstly, the maximum training error and testing error of the BP neural network are $8.8874 \times 10^{-4}$ and $1.4886 \times 10^{-3}$. This exposes another defect of BP in that its generalization ability is limited while the training process seems successful; secondly, the RBF network lacks searching ability as the maximum error obtained by it reaches $1.4557 \times 10^{-3}$ in the learning period; thirdly, the ELM approach does not perform as well as in the temperature compensation problem when the sample scale and model complexity increase; fourthly, although the maximum error of the PSO-LSSVM associated with the testing set is the best, the optimization method which is the bottleneck to sufficiently explore the parameter space for the maximum error in the training set reaches $1.2025 \times 10^{-3}$; finally, in combination with the advantages of regularization theory and kernel trick, the maximum errors of the proposed KELM method in the training set and testing set are $8.2560 \times 10^{-4}$ and $1.0836 \times 10^{-3}$, which demonstrates an ideal synthetic compensation performance at a low computation cost.

Table 6. Synthetic compensation results of training set.

| Temperature Compensation Methods | Err (min)          | Err (max)          | Err (mean)         | Err (variance)      |
|----------------------------------|--------------------|--------------------|--------------------|---------------------|
| BP                               | $7.9857 \times 10^{-7}$ | $8.8874 \times 10^{-4}$ | $1.3630 \times 10^{-4}$ | $1.8544 \times 10^{-8}$ |
| RBF                              | $3.0434 \times 10^{-7}$ | $1.4557 \times 10^{-3}$ | $1.7424 \times 10^{-4}$ | $3.3145 \times 10^{-8}$ |
| PSO-SVM                          | $5.0338 \times 10^{-7}$ | $1.2871 \times 10^{-3}$ | $2.9328 \times 10^{-4}$ | $4.5771 \times 10^{-8}$ |
| PSO-LSSVM                        | $6.2260 \times 10^{-7}$ | $1.2025 \times 10^{-3}$ | $1.4617 \times 10^{-4}$ | $2.6931 \times 10^{-8}$ |
| ELM                              | $7.1578 \times 10^{-7}$ | $2.1745 \times 10^{-3}$ | $2.4851 \times 10^{-4}$ | $1.5724 \times 10^{-7}$ |
| CSA-simplex-KELM                 | $1.4895 \times 10^{-7}$ | $8.2560 \times 10^{-4}$ | $1.3599 \times 10^{-4}$ | $1.5597 \times 10^{-8}$ |

Table 7. Synthetic compensation results of testing set.

| Temperature Compensation Methods | Err (min)          | Err (max)          | Err (mean)         | Err (variance)      |
|----------------------------------|--------------------|--------------------|--------------------|---------------------|
| BP                               | $3.5195 \times 10^{-7}$ | $1.4886 \times 10^{-3}$ | $1.2296 \times 10^{-4}$ | $1.4083 \times 10^{-8}$ |
| RBF                              | $1.9854 \times 10^{-8}$ | $1.2708 \times 10^{-3}$ | $1.3248 \times 10^{-4}$ | $1.6446 \times 10^{-8}$ |
| PSO-SVM                          | $4.7972 \times 10^{-8}$ | $1.2538 \times 10^{-3}$ | $2.7089 \times 10^{-4}$ | $3.5158 \times 10^{-8}$ |
| PSO-LSSVM                        | $1.5621 \times 10^{-7}$ | $9.6256 \times 10^{-4}$ | $9.5755 \times 10^{-5}$ | $1.1373 \times 10^{-8}$ |
| ELM                              | $4.6555 \times 10^{-9}$ | $1.8777 \times 10^{-3}$ | $1.1114 \times 10^{-4}$ | $5.3387 \times 10^{-8}$ |
| CSA-simplex-KELM                 | $3.2846 \times 10^{-7}$ | $1.0836 \times 10^{-3}$ | $1.1019 \times 10^{-4}$ | $1.0429 \times 10^{-8}$ |

All the synthetic compensation details are illustrated in Figure 9, where all the maximum synthetic compensation errors occur at $−20$ °C except for the ELM, which is due to two facts: the stress between the silicon cup and the glass becomes higher with the decrease of the environmental temperature; the ELM model is not suitable for learning the structure contained in this dataset. Since the training set
lacks information about the sensor performance without static effect, the compensation error of RBF and PSO-LSSVM in all temperature states without static pressure is the maximum. It also indicates that the generalization ability of RBF and PSO-LSSVM is not satisfactory for this problem in some way. However, the KELM alleviates both the temperature effect and the static pressure effect and achieves a robust and accurate compensation model which fulfills the engineering requirements [30].

Figure 9. Cont.
Figure 9. Synthetic compensation results obtained by different models. (a) All synthetic compensation results at −20 °C; (b) All synthetic compensation results at 0 °C; (c) All synthetic compensation results at 20 °C; (d) All synthetic compensation results at 50 °C; (e) All synthetic compensation results at 70 °C.

6. Conclusions

A synthetic compensation approach that incorporates optimization methods into the framework of KELM is presented in this research. Making full use of regularization theory and kernel trick, the kernel version of the ELM (KELM) is adopted to guarantee the learning and generalization ability of the compensation model. The CSA and simplex search are applied to perform the search of an appropriate parameter set for KELM. The proposed synthetic compensation scheme avoids a complex network topology and speeds up the computation process. The temperature compensation and synthetic compensation results show that it is an effective and efficient compensation method even when dealing with distinct datasets.

Acknowledgments: This research work is supported by the National High Technology Research and Development Program of China (“863” Program), Program No. 2014AA042001. The authors would like to express their thanks to all the editors and anonymous reviewers for improving this article.

Author Contributions: Ji Li and Guoqing Hu conceived and designed the experiments; Ji Li and Wei Peng performed the experiments; Yonghong Zhou and Chong Zou analyzed the data; Ji Li wrote the paper; Jahangir Alam SM revised the English language and technical coherence.

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

References

1. Barlian, A.A.; Park, W.-T.; Mallon, J.R., Jr.; Rastegar, A.J.; Pruitt, B.L. Review: Semiconductor Piezoresistance for Microsystems. Proc. IEEE 2009, 97, 513–552. [CrossRef] [PubMed]
2. Hoang-Phuong, P.; Kozeki, T.; Dinh, T.; Fujii, T.; Qamar, A.; Zhu, Y.; Namazu, T.; Nam-Trung, N.; Dao, D.V. Piezoresistive effect of p-type silicon nanowires fabricated by a top-down process using FIB implantation and wet etching. RSC. Adv. 2015, 5, 82121–82126.

3. Doll, J.C.; Corbin, E.A.; King, W.P.; Pruitt, B.L. Self-heating in piezoresistive cantilevers. Appl. Phys. Lett. 2011, 98, 834. [CrossRef] [PubMed]

4. Maboudian, R.; Carraro, D.; Senesky, D.G.; Roper, C.S. Advances in silicon carbide science and technology at the micro- and nanoscales. J. Vac. Sci. Technol. A 2013, 31. [CrossRef]

5. Hoang-Phuong, P.; Dzung Viet, D.; Nakamura, K.; Dimitrijev, S.; Nam-Trung, N. The Piezoresistive Effect of SiC for MEMS Sensors at High Temperatures: A Review. J. Microelectromech. Syst. 2015, 24, 1663–1677.

6. Yalamarthy, A.S.; Senesky, D.G. Strain- and temperature-induced effects in AlGaN/GaN high electron mobility transistors. Semicond. Sci. Technol. 2016, 31, 035024. [CrossRef]

7. Gakkestad, J.; Ohlickers, P.; Halbo, L. Effects of process variations in a CMOS circuit for temperature compensation of piezoresistive pressure sensors. Sens. Actuators A Phys. 2014, 293–300. [CrossRef]

8. Aryafar, M.; Hamedi, M.; Ganjeh, M.M. A novel temperature compensated piezoresistive pressure sensor. Measurement 2015, 63, 25–29. [CrossRef]

9. Yao, Z.; Liang, T.; Jia, P.; Hong, Y.; Qi, L.; Lei, C.; Zhang, B.; Li, W.; Zhang, D.; Xiong, J. Passive resistor temperature compensation for a high-temperature piezoresistive pressure sensor. Sensors 2016, 16, 1142. [CrossRef] [PubMed]

10. Futane, N.P.; Roychowdhury, S.; Roychauhduri, C.; Saha, H. Analog ASIC for improved temperature drift compensation of a high sensitive porous silicon pressure sensor. Analog Integr. Circuits Signal Process. 2011, 67, 383–393. [CrossRef]

11. Hao, X.; Maenaka, K.; Takao, H.; Higuchi, K. An analytical thermal-structural model of a gas-sealed capacitive pressure sensor with a mechanical temperature compensation structure. Sens. Actuators A Phys. 2014, 205, 92–102. [CrossRef]

12. Mozek, M.; Vrtacnik, D.; Resnik, D.; Pecar, B.; Amon, S. Compensation and Signal Conditioning of Capacitive Pressure Sensors. Inform. Midem 2011, 41, 272–278.

13. Luo, Y.; Yang, K.; Shi, Y.B.; Shang, C.X. Research of radiosonde humidity sensor with temperature compensation function and experimental verification. Sens. Actuators A Phys. 2014, 218, 49–59. [CrossRef]

14. Chae, C.S.; Kwon, J.H.; Kim, Y.H. A Study of Compensation for Temporal and Spatial Physical Temperature Variation in Total Power Radiometers. IEEE Sens. J. 2012, 12, 2306–2312. [CrossRef]

15. Fan, S.; Zhang, Q.; Qin, J. Temperature compensation of pressure sensor based on the interpolation of splines. J. Beijing Univ. Aeronaut. Astronaut. 2006, 32, 684–686.

16. Wang, H.R.; Zhang, W.; You, L.D.; Yuan, G.Y.; Zhao, Y.L.; Jiang, Z.D. Back propagation neural network model for temperature and humidity compensation of a non dispersive infrared methane sensor. Instrum. Sci. Technol. 2013, 41, 608–618. [CrossRef]

17. Ding, J.C.; Zhang, J.; Huang, W.Q.; Chen, S. Laser Gyro Temperature Compensation Using Modified RBFNN. Sensors 2014, 14, 18711–18727. [CrossRef] [PubMed]

18. Cheng, J.; Qi, B.; Chen, D.; Landry, R., Jr. Modification of an RBF ANN-Based Temperature Compensation Model of Interferometric Fiber Optical Gyroscopes. Sensors 2015, 15, 11189–11207. [CrossRef] [PubMed]

19. Moallem, P.; Abdollahi, M.A.; Mehdi Hashemi, S. Compensation of capacitive differential pressure sensor using multi layer perceptron neural network. Int. J. Smart Sens. Intell. Syst. 2015, 8, 1443–1463.

20. Qiu, H.-M.; Chen, D.; Li, G.-Y.; Huang, Y.; Chen, Z.-C.; Liang, J.-T. Temperature compensation of light addressable potentiometric sensor based on support vector machine. J. Optoelectron. Laser 2015, 26, 2272–2277.

21. Yang, J.; Mei, X.-S.; Zhao, L.; Ma, C.; Feng, B.; Shi, H. Thermal error modeling of a coordinate boring machine based on fuzzy clustering and SVM. J. Shanghai Jiaotong Univ. 2014, 48, 1175–1182.

22. Shao, J.; Liu, J.-H.; Qiao, X.-G.; Jia, Z.-A. Temperature compensation of FBG sensor based on support vector machine. J. Optoelectron. Laser 2010, 21, 803–807.

23. Vapnik, V.N. Statistical Learning Theory; John Wiley&Sons Inc.: New York, NY, USA, 1998.

24. Suykens, J.A.K.; Vandewalle, J. Least squares support vector machine classifiers. Neural Process. Lett. 1999, 9, 293–300. [CrossRef]

25. Huang, G.-B.; Wang, D.H.; Lan, Y. Extreme learning machines: A survey. Int. J. Mach. Learn. Cybern. 2011, 2, 107–122. [CrossRef]
26. Wong, P.K.; Wong, K.I.; Vong, C.M.; Cheung, C.S. Modeling and optimization of biodiesel engine performance using kernel-based extreme learning machine and cuckoo search. *Renew. Energy* **2015**, *74*, 640–647. [CrossRef]

27. Li, J.; Hu, G.; Zhou, Y.; Zou, C.; Peng, W.; Jahangir Alam, S.M. A temperature compensation method for piezo-resistive pressure sensor utilizing chaotic ions motion algorithm optimized hybrid kernel LSSVM. *Sensors* **2016**, *16*, 1707. [CrossRef] [PubMed]

28. Li, G.; Wang, F.; Xiao, G.; Wei, G.; Zhang, P.; Long, X. Temperature compensation method using readout signals of ring laser gyroscope. *Opt. Express* **2015**, *23*, 13320–13332. [CrossRef] [PubMed]

29. Xavier-de-Souza, S.; Suykens, J.A.K.; Vandewalle, J.; Bolle, D. Coupled Simulated Annealing. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2010**, *40*, 320–335. [CrossRef] [PubMed]

30. Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. Optimization By Simulated Annealing. *Science* **1983**, *220*, 671–680. [CrossRef] [PubMed]

31. Fan, S.K.S.; Zahara, E. A hybrid simplex search and particle swarm optimization for unconstrained optimization. *Eur. J. Oper. Res.* **2007**, *181*, 527–548. [CrossRef]

32. Pressure/Differential-Pressure Transmitter for Use in Industrial-Progress Measure and Control Systems—Part1:Genneral Specification. Available online: http://dbpub.cnki.net/grid2008/dbpub/detail.aspx?QueryID=31&CurRec=6&dbcode=SCF&dbname=SCSF&filename=SCSF00038855&urlid=\textdollar 9A4hF_YAuvQ5obgVAqNKPCYcEjKensW4ggI8Fm4gTk0UkAI8j8gFw (accessed on 17 April 2017).

33. Barnes, J.P.; Johnston, P.R. Application of robust generalised cross-validation to the inverse problem of electrocardiology. *Comput. Biol. Med.* **2016**, *69*, 213–225. [CrossRef] [PubMed]

34. Chang, C.C.; Lin, C.J. LIBSVM: A Library for Support Vector Machines. *Acm Trans. Intell. Syst. Technol.* **2011**, *2*, 27. [CrossRef]

35. LS-SVMLab1.8. Available online: http://www.esat.kuleuven.be/sista/lssvmlab/ (accessed on 17 April 2017).