Journal of Machine Engineering, 2019, Vol. 19, No. 4, 42–57
ISSN 1895-7595 (Print) ISSN 2391-8071 (Online)

Received: *07 January 2019* / Accepted: *02 October 2019* / Published online: *20 October 2019*

thermal effects, algorithm, machine tool, measurement

Christian NAUMANN[1*]
Matthias PUTZ[1]

A NEW MULTIGRID BASED METHOD FOR CHARACTERISTIC DIAGRAM BASED CORRECTION OF THERMO-ELASTIC DEFORMATIONS IN MACHINE TOOLS

It is a well-known problem of milling machines, that waste heat from motors, friction effects on guides, environmental variations and the milling process itself greatly affect positioning accuracy and thus production quality. An economic and energy-efficient method of correcting this thermo-elastic positioning error is to gather sensor data (temperatures, axis positions, etc.) from the machine tool and the process and to use that information to predict and correct the resulting tool center point displacement using high dimensional characteristic diagrams. The computation of these characteristic diagrams leads to very large sparse linear systems of equations which require a vast memory and computation time to solve. This is particularly problematic for complex machines and varying production conditions which require characteristic diagrams with many input variables. To solve this issue, a new multigrid based method for the computation of characteristic diagrams will be presented, tested and compared to the previously used smoothed grid regression method.

1. INTRODUCTION AND STATE OF THE ART

THERMAL EFFECTS present one of the leading causes of positioning errors in machine tools [1]. They are caused by shifting temperature distributions inside the machine tool which lead to thermo-elastic deformations. These temperature fields are shaped by a large number of heat sources and heat sinks. Important ones are waste heat from the cutting process, friction from guides and bearings, power losses from motors, coolants and the environment. Other relevant factors which influence the time-dependent temperature distribution inside a machine tool are the heat transfer coefficients and the thermal capacity which influence the rate at which heat is transferred and stored. It is this multitude of influences which makes thermal issues so difficult to handle [2].

Generally, there are two main strategies of dealing with thermal errors in machine tools: correction and compensation strategies. Correction of thermo-elastic deformations involves the prediction or measurement of temperature or deformation fields and using them to offset

[1] Fraunhofer Institute for Machine Tools and Forming Technology IWU, Chemnitz, Germany
[*] E-Mail: Christian.Naumann@iwu.fraunhofer.de
https://doi.org/10.5604/01.3001.0013.6229

and thus correct the tool center point (TCP) position. Compensation on the other hand seeks to prevent or divert the influx of heat or to prevent or direct the deformation to reduce the TCP displacement [3]. These definitions follow the conventions within the German Research Foundation project CRC/TR96 on the thermo-energetic design of machine tools and may deviate from some literature sources.

Sample correction strategies are the structure model based correction [4, 5], the indirect correction based on transfer functions [6, 7], the characteristic diagram based correction [8, 9] or the correction based on integrated deformation sensor (IDS) measurements [10].

Sample compensation strategies are active cooling [11], switchable heat storage through phase change materials (PCM) [12, 13], design optimization such as the thermo-symmetric design of machine tools [14] or the use of materials with negative thermal expansion [15].

This article focuses on the characteristic diagram based correction of thermo-elastic deformations. It uses sensor and control data as input variables to predict the TCP displacement. The underlying model for this prediction, i.e. the characteristic diagram is trained using measurement or simulation data. These contain datasets for the input variables along with their corresponding measured or simulated TCP displacements. Using this a priori trained model, fast control-internal correction algorithms can be used to calculate TCP offsets from the current sensor readings in real-time in order to improve a machine tool's positioning accuracy [16]. Characteristic diagram based correction is a special type of regression analysis and thus a black-box method requiring little to no knowledge of the machine tool or the heat transfer mechanisms therein. A similar method using multiple variables in a different multiple regression model can be found in [17].

There are several factors determining the quality of a characteristic diagram such as:

- the number and type of input variables [18],
- the type, fineness and structure of the characteristic diagram grid [19],
- the method of interpolation and extrapolation on the grid,
- the quantity and quality of the training data.

It is this third point, the interpolation and extrapolation, which will be the main focus of this paper. A generalized explanation of the computation and use of characteristic diagrams was published in [7, 20] under the synonym "Smoothed Grid Regression" (SGR). It explains that both interpolation and extrapolation are achieved through local kernel functions. There are a great number of possible kernel functions like polynomials, wavelets, radial basis functions, etc.. One of the simplest and most common functions is the piecewise multilinear kernel. The kernels are parametrized using a set of data and smoothing equations. This step, called characteristic diagram computation, involves the assembly and solution of a large, sparse linear system of equations. It is mainly this step which has limited the size and complexity of characteristic diagrams thus far.

The problem, sometimes called "curse of dimensionality", is that the grid size grows exponentially with the number of input dimensions and polynomial with the fineness of the grid. This results in extremely large linear systems which require vast amounts of working memory and computation time. Additionally, the use of a larger and finer grid with a limited amount of training data means that the characteristic diagram will be largely shaped by the smoothing conditions. This results in two major challenges:

- solving the very large linear systems in an efficient manner,

– making sure the smoothing conditions affect the solution (i.e. the characteristic diagram) in the desired way with respect to interpolation and extrapolation.

The first challenge is an old one, which has consequently many possible solutions available within the state of the art. Large linear systems (e.g. $A \cdot x = b$) have a high-dimensional coefficient matrix (A). These matrices are often so large that they can no longer be stored fully within the working memory. Since they are also most often sparse, i.e. contains mostly 0.0 entries, they can be compressed into sparse matrix formats, which store only the non-zero entries, see [21]. While significantly reducing the necessary memory demand they also complicate the matrix arithmetic. Matrix by vector multiplication usually remains easy but, e.g., matrix element assignments or retrieval become far more difficult.

By far the larger problem, however, is solving the linear system of equations. Using standard direct solvers such as the Gauß-Jordan-Algorithm requires $O(n^3)$ floating point operations (FLOPs), where n is the dimension of the matrix *A*. Since the coefficient matrix *A* is symmetric positive definite for characteristic diagrams, the Cholesky-Factorization may also be used, but it is still only slightly faster. Modern direct solvers for sparse matrices employ graph theory to efficiently traverse large sparsity patterns. An overview of such methods can be found in [21]. Beside large computation times, the matrix fill-in presents the largest issue. That means that solving the linear system directly inevitably fills many of the previously empty entries in the sparse coefficient matrix, which quickly fills up the working memory.

An alternative, therefore, present the numerical or iterative solvers. Since they only update the solution vector during every iteration and do not modify the coefficient matrix, the memory demand does not grow during the solution process. This is typically done using Krylov subspace methods such as the Conjugate Gradients (CG) method, GMRES or BiCGSTAB [22].

They improve the accuracy in every iteration so that the efficiency of these algorithms is measured in the number of cycles they need to reach a predetermined residual error threshold. This efficiency is, however, not only dependent on the matrix dimension but also on its properties, particularly its eigenvalues and eigenvectors. To improve these properties and speed-up the computation, preconditioners are often used. They modify the coefficient matrix or perform additional operations on the iterate in every cycle. Of particular interest in the context of characteristic diagrams are grid based preconditioners and solvers such as multigrid algorithms [23], multigrid preconditioned CG algorithms and BPX-preconditioned CG algorithms [24]. This idea will be discussed in greater detail in Chapter 3. An additional benefit in most iterative solvers is that (depending on the application) they can often be used without the need to ever fully assemble the coefficient matrix *A*. While this will be less efficient to compute and more difficult to implement, it eliminates the limitations on the matrix dimension almost completely and thus makes memory size nearly irrelevant (at least in terms of memory).

Smoothness of the solution or the characteristic diagram is an important requirement in most applications. This obviously requires the solution to be continuous. Depending on the application, the concept of smoothness can have different meanings, such as:

– the derivative must be continuous AND/OR,
– the second derivative must not exceed a certain threshold AND/OR,

– the derivative must jump as little as possible AND/OR,
– the derivative must not exceed a certain threshold, etc..

In the most commonly used type of characteristic diagram with multilinear kernel functions, the first two become irrelevant but the latter two are usually important for the acceptance of the solution in real-world applications. In thermal error prediction, e.g., it is usually expected that for any thermal state of the machine tool (as represented by the characteristic diagram), small temperature changes will cause a nearly linear expansion and thus a more or less linear change in the output variable. Good temperature sensors have a resolution of about 0.1 K.

That means that all thermal inputs into a characteristic diagram can change by 0.1K instantly. When that happens, the resulting change in the predicted deformation should likewise be small, no matter which thermal state (i.e. section of the characteristic diagram) the machine tool is currently in.

When the training data is sparse compared to the number of grid elements the smoothing requirements will not only ensure that the solution is smooth and the linear system has full rank, they also determine the way the solution behaves in unpopulated grid sections and particularly near the grid boundary. Depending on the application one may, e.g., wish for the characteristic diagram to converge toward a constant, toward a straight line perhaps even toward some higher-degree polynomial. How to control smoothness and achieve the desired extrapolation behaviour in characteristic diagrams will be explained in later Chapters.

To summarize: the quality of characteristic diagram based correction depends on the type of kernel function used for interpolation and extrapolation on the grid, combined with the method of smoothing. Both determine the linear system of equations which needs to be solved in order to obtain the characteristic diagram. While small linear systems can still be solved with direct methods, larger linear systems require efficient (preconditioned) iterative solvers.

Computing smooth characteristic diagrams through an FEM based approach with multigrid solvers was first suggested in [7] in 2012. In 2015 the grid fineness independent convergence was demonstrated and a study of different iterative solvers showed the multigrid preconditioned CG method to be the most efficient solver [8]. While this has already demonstrated the validity of the new approach, it has only done so for the most basic type of smoothing, the one that converges toward a constant. This article will use a simple example to show why this will often be problematic and how to modify the methods to obtain linear smoothing. To that end, Chapter 2 gives a brief overview on characteristic diagram based correction. Chapter 3 explains the new FEM based approach with multigrid solvers. Chapter 4 then uses a simple 2D example to demonstrate the deficiencies of this method. Chapter 5 shows a way to achieve polynomial extrapolation using Bogner-Fox-Schmidt-Elements, polynomial basis functions and higher-order smoothing terms. Since this variant is cumbersome for higher dimensions and can be hard to control over larger grids, another variation is suggested in Chapter 6. It uses a two-step approach to combine the smoothed grid regression with linear extrapolation on a coarse grid with the new FEM based approach on the fine grid. Chapter 7 tests the new method on machine tool thermal deformation measurement data. Finally, a summary will be given and an outlook on future research provided.

2. CHARACTERISTIC DIAGRAM BASED TCP CORRECTION

Characteristic diagrams are a fundamental tool of engineers used to approximate real-valued functions that depend on one or more input variables. The characteristic diagrams used in this paper are based on the smoothed grid regression introduced by Priber in 2003 [20] and later improved to enable efficient, high-dimensional characteristic diagrams able to approximate thermo-elastic deformations in machine tools [8]. These characteristic diagrams comprise a rectangular grid of support points and a set of kernel functions used to interpolate between them. Popular kernels are polynomials or splines, where higher-dimensional kernels are usually created by multiplying one-dimensional kernels (see [20]).

The creation of a characteristic diagram starts with the selection of input variables needed to approximate the output variable. This can, e.g. be a set of temperature variables

$$\vec{t} = (t_1, t_2, \dots, t_n),$$

where each t_i corresponds with a temperature sensor on the machine tool. The next step is to define and discretize the domain of each variable. The fineness of the discretization usually depends on the variability of the directional derivative and the type of kernel used:

$$T_1 = t_{1,\min} : dt_1 : t_{1,\max} \;;\; \dots \;;\; T_n = t_{n,\min} : dt_n : t_{n,\max}.$$

In this example of an n-dimensional equidistant discretization, each grid axis vector T_i contains the support points of the i-th temperature variable which range from $t_{i,\min}$ to $t_{i,\max}$ with a constant spacing of dt_i. The type of kernel is thus usually chosen along with the grid fineness in order to obtain optimal grids and avoid overfitting. Given a sufficiently fine grid, simple piecewise multilinear kernels are sufficiently accurate and generally well suited for the approximation of thermal deformations [8]. A simple linear one-dimensional kernel function K at temperature t is:

$$K_j(t) = \begin{cases} \frac{t_{1,j+1} - t}{t_{1,j+1} - t_{1,j}} \;, & t \in \left[t_{1,j}, t_{1,j+1}\right] \\ \frac{t - t_{1,j-1}}{t_{1,j} - t_{1,j-1}} \;, & t \in \left[t_{1,j-1}, t_{1,j}\right] \\ 0 \quad , & else \end{cases} \tag{1}$$

While complex grid structures may sometimes be useful in minimizing the necessary degrees of freedom of a characteristic diagram, simple equidistant grids often perform equally well and are usually best at avoiding overfitting in thermal error estimation [19]. The next step is the gathering of training data which comprises a set of input data and their corresponding output data. These may be obtained from measurements or simulations and should cover as much of the input domain as possible. From this training data, data fitting equations are created in a least-squares error minimization approach. The thermal error in the z-direction $z(\vec{t})$ for a set of temperatures $\vec{t}$ on a grid with N nodes might then be:

$$z(\vec{t}) \approx \sum_{i=1..N} r_i \cdot K_i(\vec{t}) =: \varphi(\vec{t}). \tag{2}$$

The thermal error is thus the weighted sum of all kernel functions evaluated at the grid location $\vec{t}$. Most kernel functions will, however, not be in the immediate vicinity of $\vec{t}$ and will

therefore be zero for any given data sample. Substituting measured or simulated temperatures and corresponding z deformations in equation (2) creates one data fitting equation per data sample. Since the data is most often sparse in comparison to the rather large grids, the assumption of smoothness is used to turn the underdetermined system into an overdetermined system by adding smoothing equations. A constant smoothing equation in 1D is:

$$r_i - r_{i-1} = 0. \tag{3}$$

A linear smoothing equation in 1D is:

$$-r_{i-1} + 2\,r_i - r_{i+1} = 0. \tag{4}$$

The resulting linear system then provides the coefficients *r* of the kernel functions for each grid vertex and thereby defines the characteristic diagram. Denoting the data fitting equations by D and the smoothing equations by S the resulting linear system is:

$$\begin{bmatrix} D \\ S \end{bmatrix} \cdot \vec{r} = \begin{bmatrix} \vec{z} \\ 0 \end{bmatrix}. \tag{5}$$

Multiplying the whole system by the transposed of the matrix DS gives a symmetric sparse linear system in the form

$$A \cdot \vec{x} = \vec{b}, \tag{6}$$

where the solution $\vec{x}$ contains the weights of the kernel functions. A detailed account of the entire smoothed grid regression (SGR) algorithm can be found in [7] and [20]. Once the model has been trained, i.e. the optimal coefficient vector $\vec{r}$ has been found, the deformations for any set of temperatures $\vec{t}$ can be predicted using eq. (2).

One possible application of characteristic diagram based interpolation is the estimation of thermal deformations from a small set of temperature sensors (strategically distributed across the machine tool) and the axis positions, which has been thoroughly investigated and tested in [7, 8, 16]. The axis positions are needed because the TCP error is usually position dependent. They can simply be added as additional input variables and treated the same as the temperature variables. Another application is the approximation of heat transfer coefficients for the accurate modelling of the heat dissipation through convection in thermal simulations of machine tools [25].

3. A NEW FEM BASED APPROACH WITH A MULTIGRID SOLVER

Chapter 1 introduced multigrid methods as efficient iterative solvers for large sparse linear systems. Unfortunately, they cannot simply be applied to any linear system. In order to work, they require a linear system created from the discretization of a partial differential equation (PDE) on a set of at least two nested (bisected) grids. The part with the nested grids is easy to facilitate for characteristic diagrams. Starting from a very coarse grid, such as a single, large rectangle in 2D or a large hypercuboid in nD, one simply bisects this grid in every grid dimension until the desired grid fineness is reached. This creates L nested grids with increasing fineness. Naturally, one does not need to bisect every dimension the same number of times, though this would require some adaptations to the algorithm.

Obtaining the linear system from a PDE instead of the combination of data and smoothing equations as described in Chapter 2 requires an entirely new approach. The construction of such a PDE for the desired data regression will now be explained. In order to be able to build on the experience gained from the smoothed grid regression (SGR), the new approach should resemble the previous algorithm as closely as possible. The SGR simultaneously minimizes the data fitting error ε^D and the smoothing error ε^S:

$$\vec{\varepsilon}^D(\vec{r}) = \vec{z} - D \cdot \vec{r},$$
$$\vec{\varepsilon}^S(\vec{r}) = \vec{0} - S \cdot \vec{r}.$$

Finding an optimal solution (set of kernel function weights $\vec{r}$) requires solving the following optimization problem:

$$\min_{\vec{r}} \ \tfrac{1}{2}\|\vec{z} - D \cdot \vec{r}\|^2 \ + \ \tfrac{1}{2}\|S \cdot \vec{r}\|^2. \tag{7}$$

Constant smoothing (see equation (3)) penalizes all non-zero gradients. Another way of doing this is by minimizing the integral over the gradients. Using this and equation (2), one can thus rewrite the optimization problem as:

$$\min_{\vec{r}} \ \tfrac{1}{2}\sum_{i=1}^{n\mathrm{Data}} w^D(\vec{t}_i) \cdot \left|z_i - \varphi(\vec{t}_i)\right|^2 \ + \ \tfrac{1}{2}\int_\Omega w^S(\vec{t}) \cdot \left|\nabla\varphi(\vec{t})\right|^2 d\Omega, \tag{8}$$

where Ω is the entire grid and $\varphi(\vec{t})$ is the characteristic diagram approximation of the z displacement. $w^D$and w^S are additional weights that can be used to assign different confidence values to data points or make local adjustments to the smoothness. In practice, w^D is almost always set to 1.0 and w^S is set to a constant value. Finding a good smoothing constant w^S depends on the range of the output variable (z) and usually requires some trial and error. In the SGR, w^S also depends on the grid fineness because it is applied to each grid vertex instead of the entire grid space Ω. For thermal error prediction, where the output variable usually has a range of 10–200 µm, good smoothing constants can most often be found between 0.1 and 1.0. While setting individual smoothing weights to control the shape of a solution in any section of the grid may sound extremely useful, it is actually very difficult to manage in large, high-dimensional grids. This is mostly because characteristic diagrams with more than two input variables can no longer be easily plotted which makes it difficult to identify grid sections where weights may need adjusting. Also, since the data is usually very sparse, the "correct" shape of the characteristic diagram can only be guessed at.

Using constant weights as suggested, the optimum is found by setting all directional derivatives of the target function in (8) equal to zero:

$$\sum_{i=1}^{n\mathrm{Data}}(z_i - \varphi(\vec{t}_i)) \cdot \delta\varphi(\vec{t}_i) \ + \ \int_\Omega w^S \cdot \nabla\varphi(\vec{t}) \cdot \nabla\delta\varphi(\vec{t})\, d\Omega = \ 0, \ \ \forall\delta\varphi. \tag{9}$$

Equation (9) is now already the weak formulation of a PDE. Since the finite element method (FEM) operates on an element-by-element basis, it would be nice to have the data fitting term also as an integral over the domain Ω. This can be accomplished using the Dirac-Function:

$$\delta_{\vec{t}_i}(\vec{t}) := \begin{cases} \infty, & \vec{t} = \vec{t}_i \\ 0, & else \end{cases}, \ \text{with} \ \int_\Omega \delta_{\vec{t}_i}(\vec{t}) \cdot \varphi(\vec{t}) d\Omega = \ \varphi(\vec{t}_i)\,. \tag{10}$$

With this Dirac-Function, one can now rewrite equation (9) as:

$$\sum_{i=1}^{n\mathrm{Data}} \int_{\Omega}(z_i - \varphi(\vec{t}_i)) \cdot \delta\varphi(\vec{t}_i) \cdot \delta_{\vec{t}_i}(\vec{t})\, d\Omega + \int_{\Omega} w^S \cdot \nabla\varphi(\vec{t}) \cdot \nabla\delta\varphi(\vec{t})\, d\Omega = 0, \quad \forall\delta\varphi. \quad (11)$$

The FEM now divides the integral over Ω into the sum of the integrals over each grid element and replaces $\varphi(\vec{t})$ and $\delta\varphi(\vec{t})$ by so-called test functions. As test functions, the same kernel functions as before (compare eq. (1)) were chosen. After computing (11) for every grid element, the coefficient matrix *A* and the right-hand-side b are assembled to form a linear system nearly identical to equation (6), which can now be solved more efficiently by multigrid methods.

In [16], this new FEM based algorithm was successfully tested and has demonstrated that it enables characteristic diagrams with as many as ten or more input variables [8]. In another test of this new method, grid fineness independent convergence was demonstrated and a study of different iterative solvers showed the multigrid preconditioned CG method to be the most efficient solver [8]. The multigrid algorithm achieves this fast convergence behaviour by using a smoothed approximation of the residual ($\vec{b} - A \cdot \vec{x}$) from the coarser grids to improve the fine grid solution during every cycle of the CG algorithm. Different types of smoothing, such as linear smoothing (see eq. (4)), require different PDEs and often different test functions. One such example will be given in Chapter 5.

4. A 2D TEST CASE FOR THE EVALUATION OF SMOOTHNESS AND EXTRAPOLATION BEHAVIOUR

The previous chapters have argued that very large characteristic diagrams require iterative solvers due to limited working memory and have demonstrated how characteristic diagram computation can be modified to enable the usage of multigrid solvers. So far this has only been done for the simplest type of smoothing which converges toward a constant. This chapter will use a simple two-dimensional function to show the limitations of constant smoothing and the need for more complex smoothing strategies. The test function is:

$$z(x,y) = \tfrac{1}{10}(x-5)^2 + \tfrac{1}{2}x + \tfrac{1}{2}y + \tfrac{1}{8}e^{x-6} - 2\,,\ x \in [0{,}10], y \in [0{,}10] \quad (12)$$

This function was designed specifically for the tests in this and the following chapters because it behaves differently in different grid sections and requires constant, linear and super-linear extrapolation in different sections of the border. Performing these tests with actual measurement or simulation data of the thermal error from some machine tool may have also been possible but a 2D example was desired for better visual evaluation and thermal error prediction with only two input variables will usually not work for realistic thermal load cases.

Since the test is supposed to focus on smoothing and extrapolation, the training data was created randomly within the subdomain $x \in [1.5{,}8.5], y \in [1.5{,}8.5]$. 500 training data points were used and each output value received a random offset between –0.5 and +0.5 to represent stochastic measurement errors and thus increase the difficulty. The tested characteristic diagram variants will be judged by their (optically perceived) smoothness and whether or not they can reproduce the qualitatively correct extrapolation behaviour.

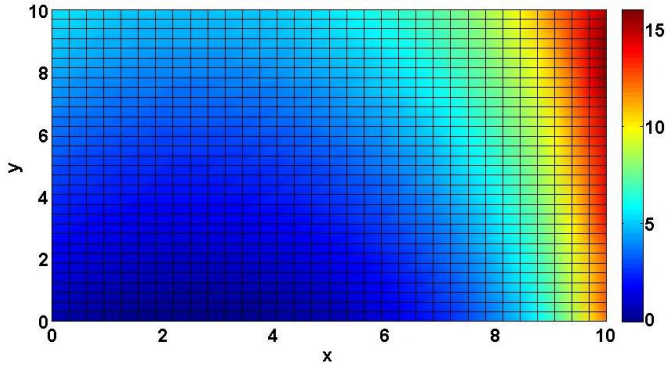

Fig. 1. The 2D test function z (x,y)

Naturally they must also reproduce the function accurately within the training data subdomain. Quantitatively exact estimations of the values on the border are not expected due to the exponential term within eq. (12) and the negative effects of the random offsets. The first test will be the SGR method with linear smoothing.

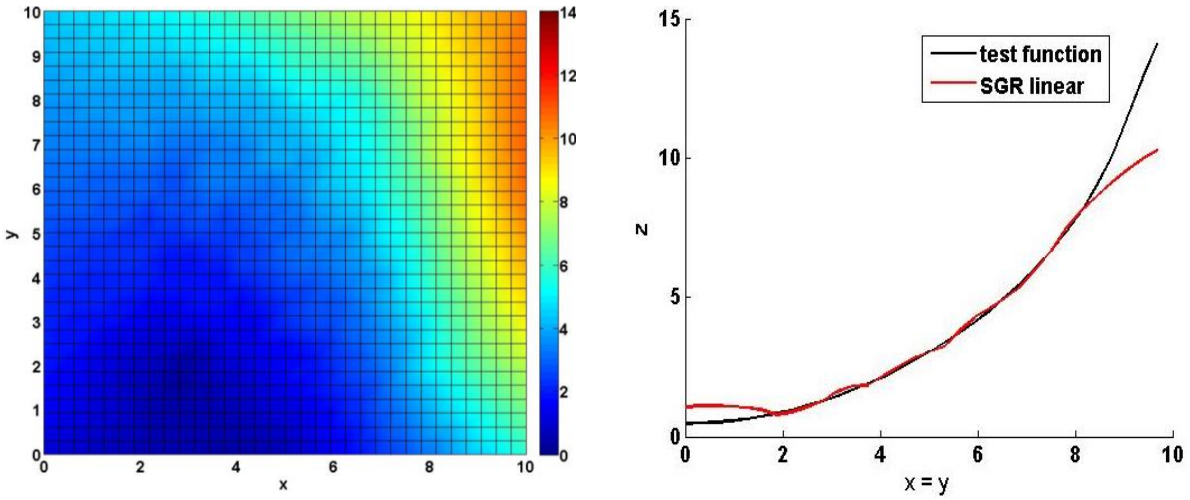

Fig. 2. SGR approximation of test function with linear smoothing conditions; left: surface plot, right: diagonal

Figure 2 clearly shows linear extrapolation beyond the training data. As expected, it could not accurately reproduce the exponential growth towards the right border. Otherwise it presents a quite good representation and extrapolation of the training data. Some similar tests with significantly less data points have shown problems in the corners. In some of these cases the corners folded in the opposite direction of where the test function lies. This was an unfortunate combination of the random data offset and the smoothing requirements. Since the corners are furthest from all other data points in this example, the closest data point will have a very strong influence on the shape of the entire section.

The second test is with constant smoothing using the FEM based approach, see eq. (11). Fig. 3 shows the expected constant extrapolation which is entirely unsuited to the test function (12).

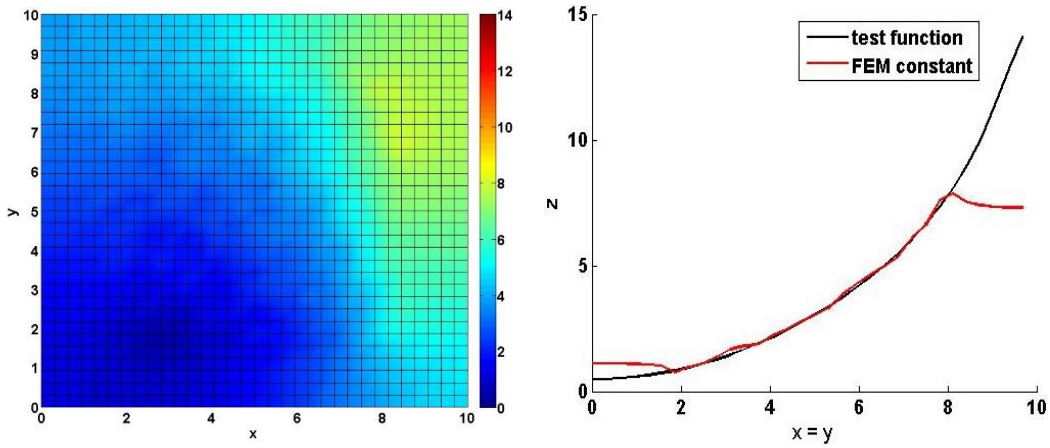

Fig. 3. Approximation of test function with the new FEM based method

It must also be noted that along the top right border the extrapolation lies below the outer most test data, see also Fig. 4. This is due to the global nature of characteristic diagrams and the constant smoothing conditions, i.e. since most of the data points are well below the few larger points in the top and right sections, the extrapolation there will be "dragged down" in order to minimize the overall smoothing error.

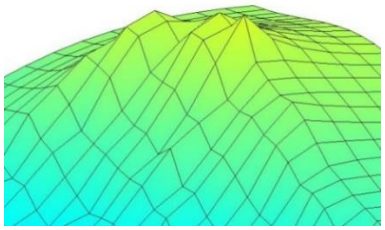

Fig. 4. 3D view of top right corner of approximation of test function with the new FEM based method

Consequently, particularly large values there will result in sharp spikes in the grid. Local manipulations of the smoothing weights might remove the spikes and improve extrapolation but it will never be linear or even super-linear. In thermal error prediction where linear thermal expansion is expected, the constant smoothing conditions will similarly fall short.

This simple example demonstrates that linear extrapolation is essential for many data fitting problems. Since there is not yet an efficient way to compute such solutions on large grids, the following two chapters will present some new methods to enable linear extrapolation with multigrid solvers.

5. HIGHER ORDER SMOOTHING WITH BOGNER-FOX-SCHMIDT ELEMENTS

Higher order smoothing, particularly linear smoothing, requires a new PDE. Therefore, we will proceed similar to Chapter 3, starting with equation (7) which still holds true. The difference now, is that the smoothing matrix S looks different and therefore the second term in equation (8) must also be changed.

$$\min_{\vec{r}} \tfrac{1}{2}\sum_{i=1}^{n\mathrm{Data}} w^D(\vec{t}_i)\cdot\left|z_i-\varphi(\vec{t}_i)\right|^2+\tfrac{1}{2}\int_\Omega w^{S_1}(\vec{t})\cdot\left|\nabla\varphi(\vec{t})\right|^2+w^{S_2}(\vec{t})\cdot\left|\Delta\varphi(\vec{t})\right|^2 d\Omega. \quad (13)$$

This new target function now contains an additional smoothing term which minimizes the integral over the second derivative along with the data fitting and constant smoothing. In analogy to Chapter 3, the weak formulation of the PDE can be obtained by setting the directional derivatives to zero:

$$\sum_{i=1}^{n\mathrm{Data}}\int_\Omega(z_i-\varphi(\vec{t}_i))\cdot\delta\varphi(\vec{t}_i)\cdot\delta_{\vec{t}_i}(\vec{t})\,d\Omega+\int_\Omega w^{S_1}\cdot\nabla\varphi(\vec{t})\cdot\nabla\delta\varphi(\vec{t})\,d\Omega+$$
$$+\int_\Omega w^{S_2}\cdot\Delta\varphi(\vec{t})\cdot\Delta\delta\varphi(\vec{t})\,d\Omega=0,\quad\forall\delta\varphi. \quad (14)$$

The problem now is that the simple linear kernel functions from before no longer work here because the test function has to be continuously differentiable. This dilemma can be resolved using Bogner-Fox-Schmidt (BFS) elements:

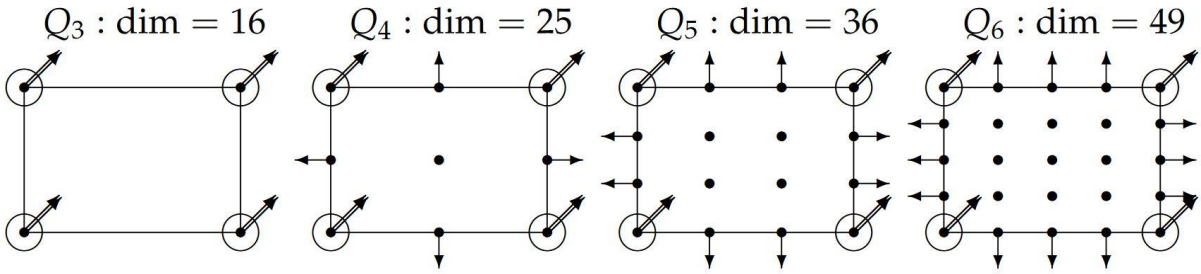

Fig. 5. The Family of Bogner-Fox-Schmidt elements in 2D [26]

In BFS elements the vertices not only hold the function values but also all directional derivatives and all mixed derivatives. In 2D that means four parameters per grid vertex instead of just one. The test functions typically used with BFS elements are from the function space Q_3 of third degree mixed polynomials (e.g. $Q_{1,2D} = \mathrm{span}(1, x, y, xy)$).

Before the new method is applied to the test function from chapter 4, some important properties should be pointed out. The ability of higher-degree extrapolation is obvious although strictly linear extrapolation may be difficult with Q_3 ansatz functions. The greatest weakness of this new method is the much greater number of parameters. Particularly with higher-dimensional grids, this not only increases the matrix size dramatically, it also worsens the ratio of data to degrees of freedom and it makes the implementation significantly more complicated. Mitigating these downsides, the new method also has two major advantages over the standard SGR, namely the fact that the solution can be computed without needing to assemble the coefficient matrix (using FEM and iterative solvers) and that the additional degrees of freedom allow the reproduction of more complex functions with coarser grids.

Figure 6 shows the approximation of the test function with a single grid element. It already gives a quite good approximation of the actual function and clearly demonstrates both the smoothness and the desired extrapolation behaviour.

Figure 7 shows the approximation with a much finer grid, where the method has clearly failed. Aside from the very large number of degrees of freedom, especially in less populated grid sections, the problem lies with the higher-degree polynomial ansatz functions.

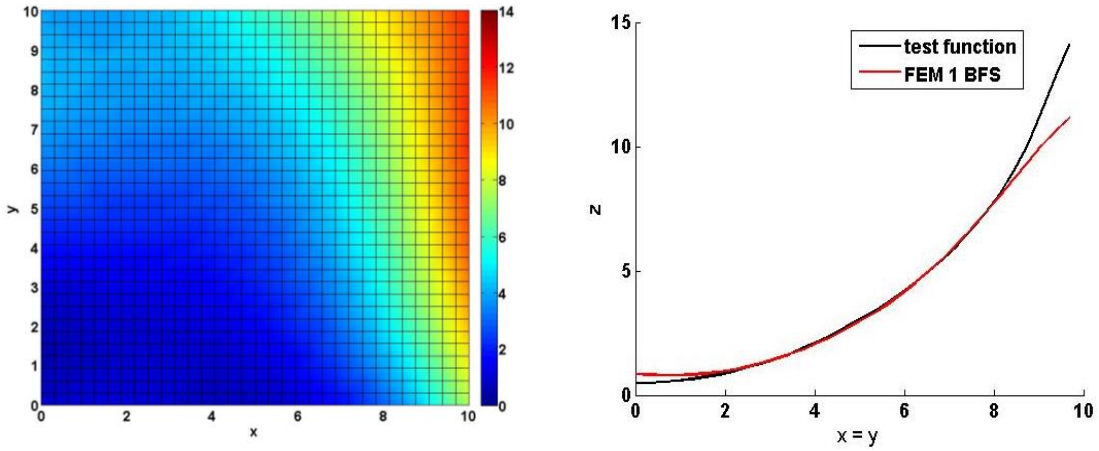

Fig. 6. Approximation of test function with the new FEM based method and one BFS element

A closer examination revealed, that they can form symmetrical wave shapes on the grid elements which still have minimal or even zero smoothing error according to equation (13). Local adjustments of the smoothing weights will likely have little effect on this. In order to best use the new method, one should therefore keep the grid so coarse that most or ideally all grid elements contain data points. This, of course, limits its usefulness.

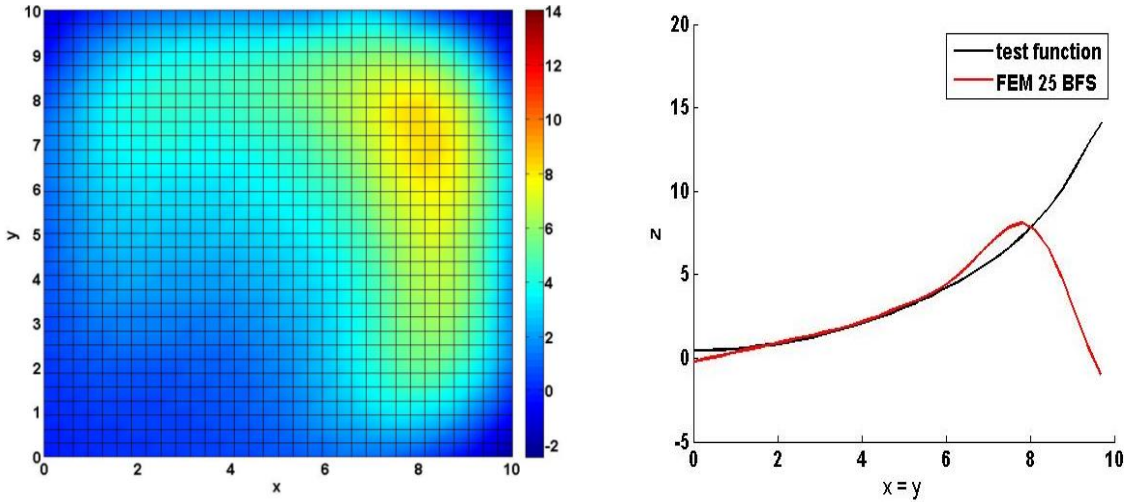

Fig. 7. Approximation of test function with new FEM based method and 25 BFS elements

6. TWO-STEP COMBINATION OF SGR AND FEM BASED APPROACH

Since the higher-order polynomials from Chapter 5 have not fully satisfied the need for a reliable linear smoothing and extrapolation method, another way must be found. For this, a short recapitulation of some important points from the previous chapters will now be given. The SGR method is well suited for data fitting problems and capable of linear extrapolation. This is, however, only possible up to a certain grid size, so that for a given number of input variables the achievable grid fineness is ultimately limited. The new FEM based approach

generally interpolates well and can handle almost any grid size. It is, however, not capable of linear extrapolation and can sometimes become spiky when the data density becomes too low.

Examining this last property more closely for different grid configurations, one will find, that the spikes appear, when isolated data points near the outskirts of the grid have several more grid elements but no more data points between them and the border. In these cases, any gradient the characteristic diagram may have had up to that data point will be smoothed out and a spike will appear depending on the data and smoothing weights. Figure 4 illustrates this effect. On the other hand, so long as all grid elements, even if they contain no data points themselves, are surrounded by some data points in every direction, no matter how far away they may be, the interpolation works just fine.

The new method combines the strengths of both the SGR and the FEM based approach by computing the characteristic diagram in two steps. The first step uses the SGR with linear smoothing (see eq. (4)–(6), [7]) on a grid that is coarse enough to still be computed with a direct solver. While for higher dimensions, this solution will likely not be sufficiently accurate, it still produces good extrapolations towards the grid boundaries. Therefore, this coarse characteristic diagram can be used to compute the function values along the borders for a much finer grid. These values are then appended to the training data. This extended set of training data can then be fitted with the FEM based approach on the finer grid (see eq. (11)). Because all grid elements are now surrounded by plenty of data points in all directions, the problems with smoothing and extrapolation are solved. Figure 8 shows the approximation of the test function with this new combination approach.

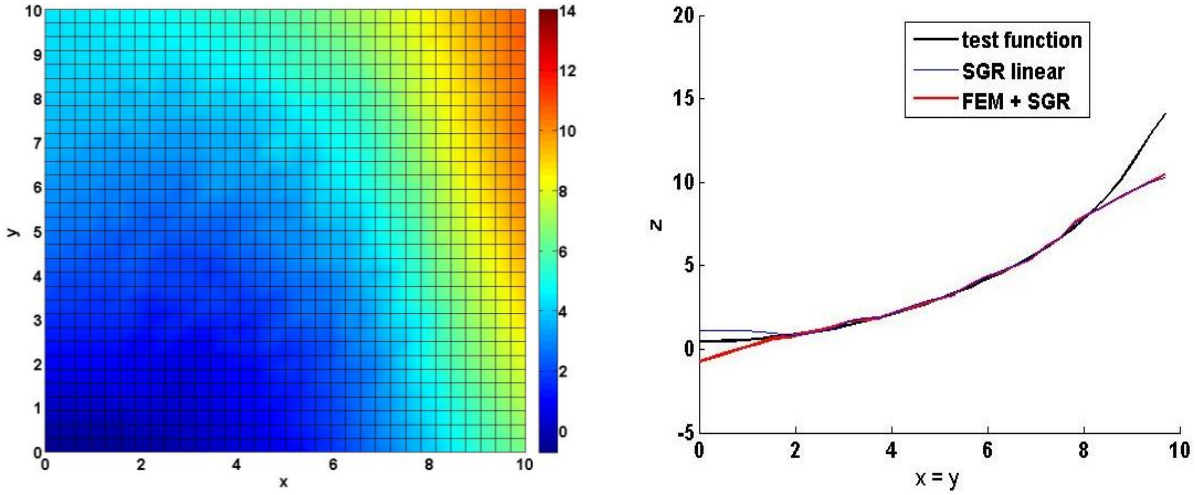

Fig. 8. Approximation of test function with combined (SGR+FEM) method

7. VALIDATION ON MACHINE TOOL MEASUREMENT DATA

In order to show that the new combination approach from the previous chapter also works for higher dimensions (i.e. more input variables) and for thermo-elastic deformation prediction, a 5-axis milling machine was measured under varying dry-cutting conditions and

relatively constant ambient conditions over several hours. Figure 9 shows a comparison of the measured z displacement of the machine tool's TCP with the estimations from both the old FE based method with constant smoothing (see eq. (11)) and the new combination approach. Since only a few randomly chosen data points were used for model training, the prediction models require both good interpolation and extrapolation capabilities. For this test, four temperature sensors were chosen. Figure 9 shows that the new combination not only works well in higher dimensions but it also outperforms the simple FE based method with constant smoothing.

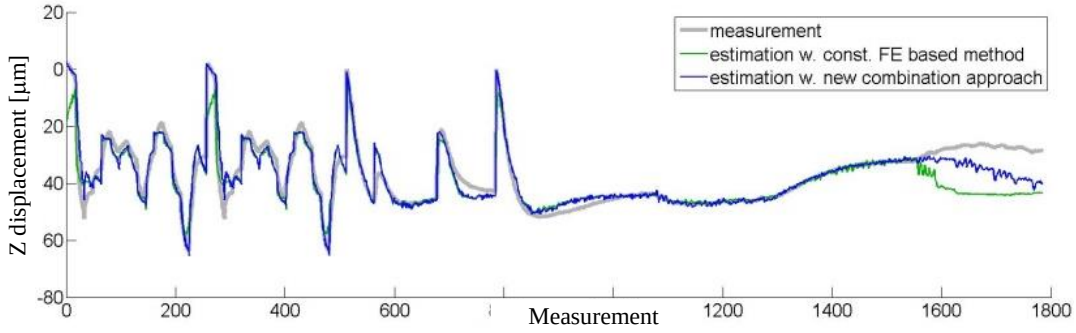

Fig. 9. Comparison of measured and estimated z displacement from a machine tool

8. SUMMARY, CONCLUSION AND OUTLOOK

Thermal effects in machine tools continue to be one of the major sources of positioning errors in machine tools. Different compensation and correction strategies can be used to mitigate these thermal errors. Characteristic diagram based correction presents a data-driven approach which maps a set of input variables directly onto the tool center point displacement. The achievable quality of this error estimation depends, among other factors, on the mechanism of interpolation and extrapolation. For characteristic diagrams this is determined mainly by the kernel functions and the smoothing conditions and weights used in the least-squares optimization. The original Smoothed Grid Regression (SGR) method for calculating characteristic diagrams can be used with constant and linear smoothing conditions, but has limitations in the grid size due to its use of direct linear system solvers. A new FEM based approach with iterative multigrid solvers which mimics the SGR with constant smoothing has been presented and has overcome the grid size limitations. A similar method for linear smoothing which is required for linear extrapolation and necessary in many real-world applications had, however, been missing. In an adaptation of the FEM based approach for constant smoothing, a new method has been constructed for polynomial extrapolation. It relies on Bogner-Fox-Schmidt elements and third degree polynomial ansatz functions. The new method has achieved polynomial extrapolation and decent approximations for coarse grids but is not well suited for fine grids. Another approach has used a combination of the SGR with linear smoothing on a coarse grid with a fine grid approximation using the FEM based algorithm. Therein the SGR computes the values on the boundaries, which are then added to the training data for the FEM based method. This method has shown good

interpolation and linear extrapolation similar to the results of the SGR with linear smoothing on the same fine grid. All tests were performed in 2D on an artificially constructed sample function. Despite this, the results are expected to transfer to most other functions and also to higher dimensional input spaces. To demonstrate the effectiveness of the new method on actual thermal deformation data, a machine tool was measured under varying thermal loads and the resulting *z*-deformation was estimated. In this test the new combination approach quite accurately estimated the thermal displacement and it outperformed the FE based approach with constant smoothing.

In Chapter 1, the quantity and quality of the training data was mentioned as another important factor for achieving optimal approximation results. Future research will therefore be focused on how to best obtain good and enough training data from thermo-elastic simulations or measurements with the least effort. This is particularly relevant for the industrial application of characteristic diagram based correction because on the one hand every additional measurement or simulation costs money and on the other hand the machine tool manufacturers must be certain that their correction algorithms are prepared for all conceivable thermal deformation states of their machine tool.

ACKNOWLEDGMENTS

This research was funded by a German Research Foundation (DFG) grant within the Collaborative Research Centers (CRC) / Transregio 96, Project ID 174223256 – TRR96 which is gratefully acknowledged.

REFERENCES

[1] BONSE R., McKEOWN P., WECK M., 1995, *Reduction and compensation of thermal errors in machine tools*, Annals of the CIRP, 44/2, 589–598.
[2] BRIAN J., 1990, *International status on thermal error research*, Annals of the CIRP, 39/2, 645–656.
[3] GROSSMANN K., et al., 2015, *Thermo-Energetic Design of Machine Tools*, Springer, 1–11.
[4] JEDRZEJEWSKI J., MODRZYCKI W., 1992, *A New Approach to Modelling Thermal Behaviour of a Machine Tool under Service Conditions*, CIRP Annals, 41/1, 455–458.
[5] ESS M., 2012, *Simulation and Compensation of Thermal Errors of Machine Tools*, Dissertation ETH Zurich.
[6] MORIWAKI T., SHAMOTO E., 1998, *Analysis of Thermal Deformation of an Ultraprecision Air Spindle System*, CIRP Annals, 41/1, 315–319.
[7] BRECHER C., HIRSCH P., WECK M., 2004, *Compensation of thermo-elastic machine tool deformation based on control internal data*, CIRP Annals Manufacturing Technology, 53/1, 299–304.
[8] NAUMANN C., PRIBER U., 2012, *Modellierung des Thermo-Elastischen Verhaltens von Werkzeugmaschinen mittels Hochdimensionaler Kennfelder*, Proceedings Workshop Computational Intelligence, Dortmund, Germany.
[9] IHLENFELDT S., NAUMANN C., PRIBER U., RIEDEL I., 2015, *Characteristic Diagram Based Correction Algorithms for the Thermo-elastic Deformation of Machine Tools*, Proceedings 48th CIRP CMS, Naples.
[10] BRECHER C., KLATTE M., WENZEL C., 2015, *Application of Machine Integrated Deformation Sensors*, Proceedings of the 11th International LAMDAMAP Conference, Huddersfield. UK, 8–17.
[11] DENKENA B., et al., 2011. *Effiziente Fluidtechnik für Werkzeugmaschinen*, Wt Werkstattstechnik online, 101/5, 347–352.
[12] DROSSEL W.G., VOIGT I., 2016. *Wärmespeicher in Werkzeugmaschinen*, VDI-Z, 158/12, 42–44.
[13] DROSSEL W.G., LAUER M., SCHNEIDER D., VOIGT I., 2016, *Development and examination of switchable heat pipes, Applied thermal engineering*, 99, 857–865.
[14] GIM T., et al., 2001, *Ball screw as thermal error compensator*, Proceedings of the 16th ASPE Annual Meeting.
[15] MARCKS P., UHLMANN E., 2008, *Compensation of Thermal Deformations at Machine Tools using Adaptronic CRP-Structures*, Proceedings 41st CIRP Conference on Manufacturing Systems, Springer, 183–186.

[16] HERZOG R., et al., 2015, *Correction Algorithms and High-Dimensional Characteristic Diagrams*, Thermo-energetic Design of Machine Tools, Lecture Notes in Production Engineering, Springer, 159–174.
[17] LI Y., et al., 2014, *Thermal error modelling of the spindle based on multiple variables for the precision machine tool*, International Journal of Advanced Manufacturing Technology, 72, 1415–1427.
[18] NAUMANN C., IHLENFELDT S., PUTZ M., 2018, *On the Selection and Assessment of Input Variables for the Characteristic Diagram Based Correction of Thermo-Elastic Deformations in Machine Tools*, Journal of Machine Engineering, 18/4, 25–38.
[19] NAUMANN C., et al., 2017. *Optimized Grid Structures for the Characteristic Diagram Based Estimation of Thermo-elastic Tool Center Point Displacements in Machine Tools*, Journal of Machine Engineering, 17/3, 36–50.
[20] PRIBER U., 2003. *Smoothed Grid Regression*, Proceedings Workshop Fuzzy Systems, Dortmund, Germany, 13, 159–172.
[21] DAVIS T.A., RAJAMANICKAM S., SID-LAKHDAR W.M., 2016, *A survey of direct methods for sparse linear systems*, Technical Report, Texas A&M University.
[22] VORST H.A., 2002, *Efficient and reliable iterative methods for linear systems*, Journal of Computational and Applied Mathematics, 149, 251–265.
[23] HACKBUSCH W., 1985, *Multigrid Methods and Applications*, Springer.
[24] PENG J., et al., 2017, *BPX-Like Preconditioned Conjugate Gradient Solvers for Poisson Problem and Their CUDA Implementations*, Advances in Intelligent Systems and Computing, 454, 633–643.
[25] GLÄNZEL J., IHLENFELDT S., NAUMANN C., PUTZ M., 2016, *Decoupling of fluid and thermo-elastic simulations of machine tools using characteristic diagrams*, Proceedings CIRP ICME 2016, Ischia, Italy.
[26] ZHANG S., 2010, *On the Full C_1-Q_k Finite Element Spaces on Rectangles and Cuboids*, *Advances in Applied Mathematics and Mechanics*, 2/6, 701–721.