Association of Metallurgical Engineers of Serbia *Scientific paper*
AMES *UDC: 004.738.5:669 ; 001.102*

A MODEL FOR CROSS-REFERENCING AND CALCULATING SIMILARITY OF METAL ALLOYS

Svetlana Pocajt[1*], Mirjana Ristić[2], Aleksandra Perić-Grujić[2], Viktor Pocajt[2], Karlo Raić[2]

[1] *Key to Metals AG, Zrenjaninski put 13p, Belgrade, Serbia*
[2] *Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, Belgrade, Serbia*

Received 30.06.2013
Accepted 07.10.2013

Abstract

This paper presents an innovative model for the comparison and cross-referencing of metal alloys, in order to determine their interchangeability in engineering, manufacturing and material sourcing. The model uses a large alloy database and statistical approach to estimate missing composition and mechanical properties parameters and to calculate property intervals. A classification of metals and fuzzy logic are then applied to compare metal alloys.

The model and its algorithm have been implemented and tested in real-life applications. In this paper, an application of the model in finding unknown equivalent metals by comparing their compositions and mechanical properties in a very large metals database is described, and possibilities for further research and new applications are presented.

Keywords: similarity of alloys, cross-referencing, equivalent metals, metal composition, mechanical properties

Introduction

Increasing international manufacturing, engineering, trade and global sourcing have led to a strongly increased need for international metals comparison, cross-referencing and interchanging, as a way to practically handle different international metals and specifications, and to select and use them effectively in the globalized manufacturing environment. However, there are many thousands of different metal alloys and consequently it is not easy to structure metal specifications in a uniformed way. Metal specifications are vital for classifying, evaluating, and specifying the chemical, mechanical, and metallurgical properties of different types of metal alloys

[*] Corresponding author. Svetlana Pocajt, s.pocajt@keytometals.com

that are used in nearly all manufacturing activities and hence of paramount importance for designers and manufacturing engineers [1].

Metal specifications are often defined by standards, which vary from country to country and can be based on different criteria. Since most of the metal specifications are complex documents, the comparison of metals, or cross-referencing, is also a challenging task. The main factors for estimating equivalency include composition, manufacturing method, finishing method, product shape, mechanical and physical properties. Additional factors such as hardenability, corrosion and heat resistance and other properties may also have significant influence on metal equivalency and potential interchangeability. Nevertheless, equivalent metals from around the world are usually grouped on the basis of composition and, to some extent, mechanical properties [2, 3].

Cross-reference tables of equivalent metals often come from recommendations of Standards Development Organizations (SDOs). However, many SDOs do not issue these kinds of guidelines, which in turn creates great difficulties in the international engineering and manufacturing environment, comprising hundreds of thousands of metal alloys and variations.

Unfortunately, besides being incompatible, specifications for metal alloys are often ambiguous and incomplete, with many missing values and intervals. For example, in specifications and standards for metal alloys, the content of alloying elements is given in the form of intervals, which sometimes define the minimum inclusion of an alloying element, e.g. chrome in stainless steel, or maximum, e.g. sulfur content in steel, and only sometimes both [4]. The base element, which forms the main metal matrix like iron in steel, is then assumed to comprise the rest of the balance [5].

However, the fact that the content of alloying elements and metal properties are defined with such incomplete intervals is often inconvenient and creates difficulties in various fields such as alloy identification, engineering calculations and notably in comparison and international cross-referencing of metals, creating again a source of possible costly errors in international engineering, material sourcing and manufacturing.

This paper presents a new model for the estimation of missing parameters in metals composition and properties, by using a statistical approach on a very large alloy database, consisting of over 180,000 alloys [6]. As briefly described in [7], this model applies classification and then fuzzy logic to calculate similarity of a metal with a given chemical composition with other metals and to discover “unknown” equivalents, that is similar materials which have been overseen by SDOs. Such “close matches” can provide engineers with a good starting point for further investigation of material equivalency and interchangeability in the process of material selection [2].

This paper explains the principles of the algorithm and its application to real-life applications of metal cross-referencing. An evaluation, which was done by applying the model on a large database of metal alloys broadly used by engineering and manufacturing companies worldwide, will be briefly presented as well.

The Model Description

This model uses one version of SmartMetals, a general algorithm designed to search and identify similar metals from a database containing metal properties information [8]. From several possible applications of the SmartMetals algorithm, this model covers two examples:

(1) Finding unknown equivalent materials (cross-references) to a selected metal alloy, e.g. a metal alloy selected by user
(2) Identifying an unknown metal when its chemical analysis is known, e.g. obtained by spectrometer.

For these applications, the model derived from SmartMetals applies the following logic sequence:

(i) Starting from the composition of a selected metal (application (1), cross-referencing) or a composition obtained by spectrometer (application (2), material identification), the model classifies the composition using rule-based reasoning and pre-defined, built-in metal classification; depending on the class of metal selected, applies weights and threshold parameters to distinguish the importance of alloying elements.

(ii) Uses statistical methods to calculate missing minimums or maximums in chemical composition and properties.

(iii) Applies fuzzy logic-based calculation to determine the level of similarity between the initial composition given and the composition of each of the metals in the database; the initial composition is either the composition of the metal for which equivalents are being searched or an analysis obtained from spectrometer.

In this research, the focus has been on the application (1), that is cross-referencing, with a specific emphasis on step (ii), where new algorithms have been applied for calculating missing values of the alloying element's content. For the other two steps, the previously developed algorithms for material identification have been re-used [9] and modified where necessary. The three steps of the model's logic sequence are described in more detail in the next three chapters.

The Classification of Metals

SmartComp classifies metal alloys from a reference database into a number of groups, which are organized hierarchically, as shown in Fig. 1. The classification can be manually pre-defined and adapted by a researcher for each application without changing the basic algorithm.

In this case, the KEY to METALS database has been used as reference [10]. We have defined 30 classes of alloys: 8 ferrous and 22 nonferrous; the second level of hierarchy was defined only for ferrous alloys, e.g. high carbon high alloyed steel.

For each group of metals, a different set of weight and threshold parameters are assigned to alloying elements, depending on their relative importance. These weights and threshold parameters, which are used for calculating the similarity of metals, have been set by using metallurgical expertise, and fine-tuned and verified across a large number of experimental datasets of metal compositions measurements obtained from metallurgical laboratories [9]. Besides their mathematical importance for the model and its calculations, the weights and thresholds have a sound physical importance:

– The weight denotes relative importance of the alloying element from the metallurgical (e.g. forming of hard carbides) and material properties perspective (e.g. strengthening, carburization etc.).

– The threshold indicates at which level the element is considered in calculations; below this value, the element is treated as an impurity. The thresholds were set to be a boundary of spectrometer sensitivity for the corresponding alloying elements [11].

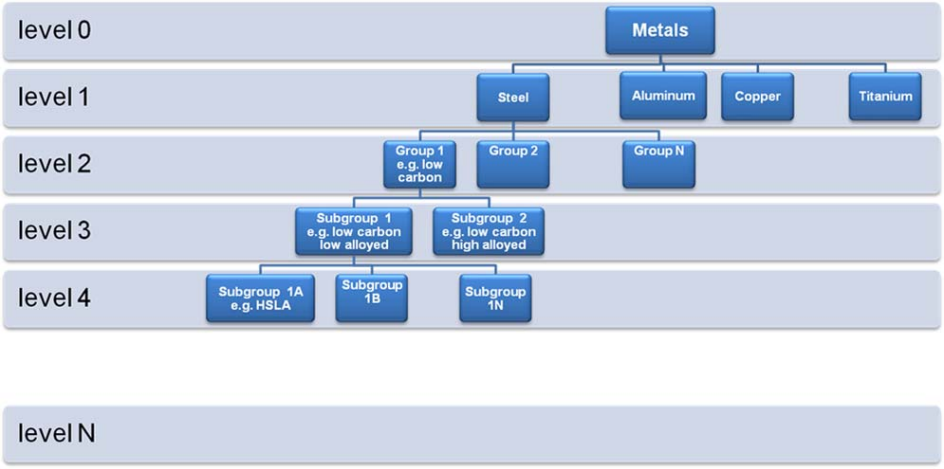

Fig. 1 An example of classification of metals in SmartMetals [8]

The Calculation of Missing Values

Since in most cases intervals for the alloying elements are not completely defined, in order to perform any meaningful comparison of compositions the system needs to infer or calculate missing minimums or maximums, thus practically forming a "virtual alloy content". Obviously, this step is critical for successful metal identification or metal comparison.

In this model, we have used a statistical approach on the KEY to METALS database for calculating missing values. Owing to large number of alloys (over 180,000) this database has always been able to provide us with the statistically significant number of alloys in each alloy group of interest. The approach will be explained on the example of calculating the values of silicon in low carbon low alloy steels.

In the first step, the database is searched to find all cases where silicon has both minimum and maximum values defined by standard specifications (Table 1).

Table 1. Minimum and maximum values of silicon content in a group of alloys as defined by standard specifications

Alloy No.	Element	min	max
1	Si	0.5	1.0
2	Si	0.7	1.0
3	Si	1.7	2.0
4	Si	1.5	3.0
....	Si		
1895	Si	0.1	0.4

Although the alloys which have the complete interval for silicon defined are a minority of the low carbon low alloy steels in the database (totaling to over 15,000), their number is still fairly high, which gives a good assumption what the alloying ranges can be.

In the next step, the interval is calculated for each record (Table 2.), as well as alloying content deviation, according to the equations:

$$Si_{int} = Si_{max} - Si_{min} \quad (1)$$

$$Si_{dev} = ((Si_{max} - Si_{min})/ Si_{max} \quad (2)$$

where: Siint is interval of the content of silicon for the alloy
Simax maximum content of silicon, as defined by the alloy specification
Simin minimum content of silicon, as defined by the alloy specification
Sidev relative deviation of the content of silicon for the alloy

Table 2. Intervals and deviations of silicon content for the group of alloys.

Alloy No.	Element	min	max	interval	Deviation
1	Si	0.5	1.0	0.5	0.50
2	Si	0.7	1.0	0.3	0.30
3	Si	1.7	2.0	0.4	0.19
4	Si	1.5	3.0	1.5	0.50
....	Si				
1895	Si	0.1	0.4	0.3	0.75

Then, the average deviation is calculated for all available alloys:

$$\overline{Si_{dev}} = \frac{\sum Si_{dev(i)}}{n} = 0.56 \quad (3)$$

where *n* is the total number of alloys in the group (1895 in this case)

Using the average deviation, the missing values are calculated in the following manner:

$$Si_{\min} = Si_{\max}(1 - \overline{Si_{dev}}) \quad (4)$$

$$Si_{\max} = \frac{Si_{\min}}{(1 - \overline{Si_{dev}})} \quad (5)$$

This procedure is then repeated for all alloying elements relevant for the group of alloys. In the case of low alloy low carbon steel, they are aluminum, boron, carbon, manganese, etc.

After that, the whole procedure is repeated again for all the groups from the database, thus forming complete virtual content tables. For a large database such as the one used in this research, the calculation tasks may take several days on an advanced PC. It is important to note that this calculation is carried out only once, as a batch

procedure; as after the virtual content for all alloys in the database is calculated, it can be stored into a database table and used for further calculations.

The Calculation of Compositional Similarity

In SmartMetals, similarity is defined as a degree of overlapping of material properties. In the material identification, this would mean overlapping of the content of each of the alloying elements (e.g. carbon, manganese and chrome) in a certain alloy X with the content of the same elements (carbon, manganese and chrome in this example) measured by spectrometer [11]. In the material cross-referencing, which is the focus of this research, the similarity applies to overlapping to the intervals of the alloying content (again e.g. carbon, manganese and chrome) in an alloy X and to the same intervals in an alloy Y. The intervals are defined by material specifications, issued in most cases by SDOs and sometimes coming from proprietary sources such as large metal producers and engineering companies.

The degree of composition overlapping can normally go from 0 (no overlapping) to 1 (the intervals of the alloying element content defined for the alloys X and Y are equal). There is also a fuzzy part, when there is a partial overlapping or when the intervals are not overlapping, but are relatively close to each other; consequently, the similarity level for the alloying element take a value between 0 and 1. For example, if the content of chromium defined by standards for alloy X is between 9 and 12% and the chromium content defined for alloy Y is between 10 and 14%, there will be a partial overlapping of the intervals. If the chromium content for alloy Z is from 13 to 15%, there will be no overlapping. However, in order to have a good comparison, the fact that these intervals are much closer than if alloy Z were to have a maximum chromium content e.g. 2% or no chromium defined as an alloying element at all, has to be taken into account. In reality, relatively smaller differences of key alloying elements can still mean that materials belong to the same group (stainless steels in this example), while large differences may indicate that materials don't belong to the same group and hence have considerably lower similarity.

This reasoning is implemented by applying fuzzy logic, the method which has already been used in this field, for ranking a fuzzy knowledge-based decision support system for materials selection in manufacturing environments [12]. For a certain alloying element, the degree of overlapping decreases linearly from 1, when intervals for the alloying element defined by standards are exactly the same (exact match) to 0, when the value measured is outside this range plus tolerance (no overlapping). In addition, when there is no overlapping on key alloying elements, the system may additionally decrease the similarity level by filtering out the alloy by applying the material classification given in Fig. 1 (that is, when it is concluded that material Y doesn't belong to the same group like X). When the measured value is outside the range defined by standards, but within the tolerance defined, the similarity level will be between 0 and 1.

The same comparison of alloying elements in the alloys Y versus the alloying elements in the reference alloy X is then repeated for all elements contained in the specification of the alloy X, excluding those filtered out by thresholds.

The total similarity factor between alloy Y and the reference alloy X is calculated as an average of similarities for all individual alloying elements, by applying weight factor for each alloying element according to the average weighted sum formula, which is very often used in fuzzy systems [13]:

$$S_t = \frac{\sum s_i w_i}{\sum w_i} \tag{6}$$

where

S_t is the total similarity between the alloy given and the reference alloy

s_i is the similarity factor for i-th alloying element

w_i is the weight (impact) factor on total similarity for the i-th element

After determining the similarity factor with alloy Y, the same procedure is repeated with all relevant materials within the database. The number of alloys that can be used is practically unlimited, for example in the KEY to METALS database, the total number of alloys is over 180,000. The results list contains alloys with the similarity factors to the reference alloy X, e.g. in descending order of similarity and above a selected similarity threshold of 0.8.

Model Verification and Applications

The international cross-referencing of metals is one of the most important applications of KEY to METALS database; its "classic" cross-reference tables which have been made using the recommendations of SDOs have been used by the global engineering community for years, Fig. 2.

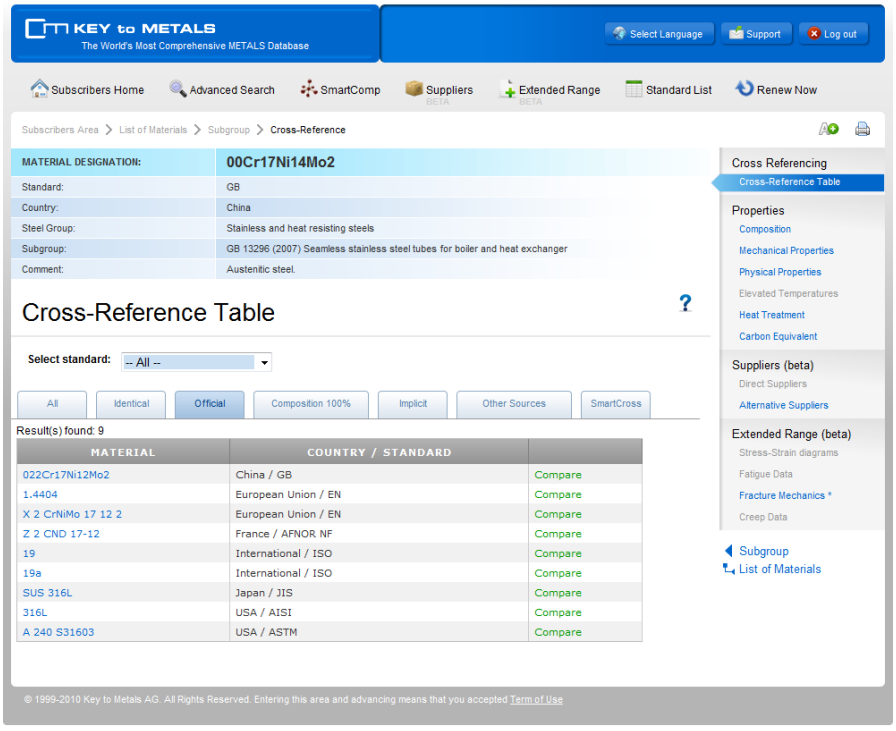

Fig. 2 An example of "classic" cross-reference table in KEY to METALS [10]; the recommendations for cross-references come from SDOs.

However, the recommendations of SDOs are intrinsically narrow and cover only a limited number of similar alloys (usually several) of the total of 180,000 alloys, thus excluding many potential equivalents. Consequently, there was a substantial practical interest to apply this model in discovering "unknown" equivalents, which means similar materials which haven't been included into the recommendations of SDOs.

After initial testing, the model has been applied to the complete KEY to METALS database. The total processing time to run all batch procedures was approximately 6 days, and the result was that the model discovered:

1) 49,676 alloys with the similarity of 100% with at least one other alloy; nearly 20% of them haven't been declared by any SDO
2) 7,958 alloys with the similarity with at least one other alloy between 95% and 100%
3) 10,074 alloys with the similarity with at least one other alloy between 90% and 95%.

These results were proved as very good, since the similarity of 70% was already consider relevant, i.e. users could consider alloys with this level of similarity as potential candidates for equivalents, depending on the application of course. Therefore, the model was implemented into the KEY to METALS cross-reference tables under the name SmartCross, thus adding the last, most complex element to the cross-referencing continuum, Fig. 3. A usual pattern of a KEY to METALS user in searching equivalent materials and candidates for equivalents is to go from more straightforward to more complex comparisons, so in many cases this model is used most complex situations, when other sources cannot provide adequate results.

An example of the implementation of the model is given in Fig. 4. The list of similar materials, which can be filtered and sorted in different ways, is given with the similarity factors. It should be noted that in this implementation it takes on average only 10 seconds to get the result list from the KEY to METALS database of over 180,000 alloys. An interesting feature is a link for side-by-side comparison, which shows the properties of the original and selected material side-by-side, thus providing the user with an opportunity to immediately check whether the differences in composition and mechanical properties are relevant for his/her particular application.

Fig. 3 Cross-referencing continuum in KEY to METALS [10]; the model is implemented as the last, most complex mean of cross-referencing.

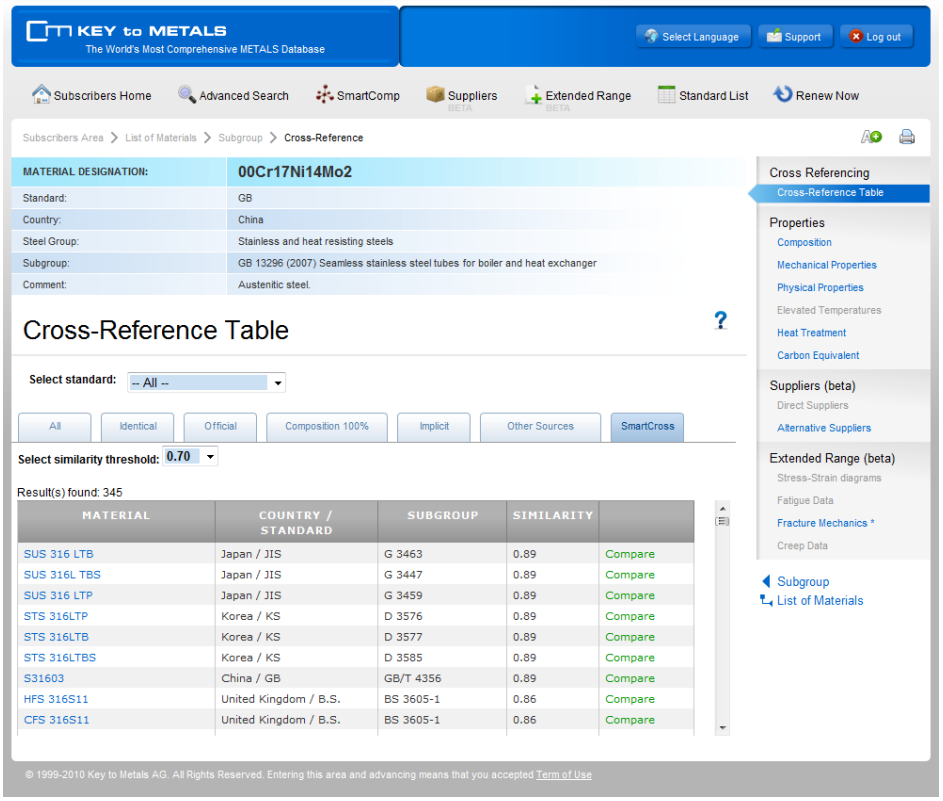

Fig. 4 An example of the implementation of the model in the KEY to METALS database.

A more recent application of the model implies extending it by adding mechanical properties to alloy comparison [6]. The calculation of missing values and similarity as given in chapters 2.2 and 2.3 is applied to the yield and tensile strength of alloys and thus mechanical similarity is obtained. This implementation also enables comparing alloys using a user-defined combination of mechanical and compositional similarity; the user can select the relative importance of mechanical vs. chemical similarity simply by moving the slider, Fig 5. The total similarity between two alloys is then calculated as a combination of chemical and mechanical similarity

$$S_t = S_c\left(1-\frac{x}{100}\right) + S_m\frac{x}{100} \tag{7}$$

where

S_t is the total similarity between the alloy given and the reference alloy
S_c is the chemical similarity between the two alloys
S_m is the mechanical similarity between the two alloys
x is the relative importance of mechanical properties in %

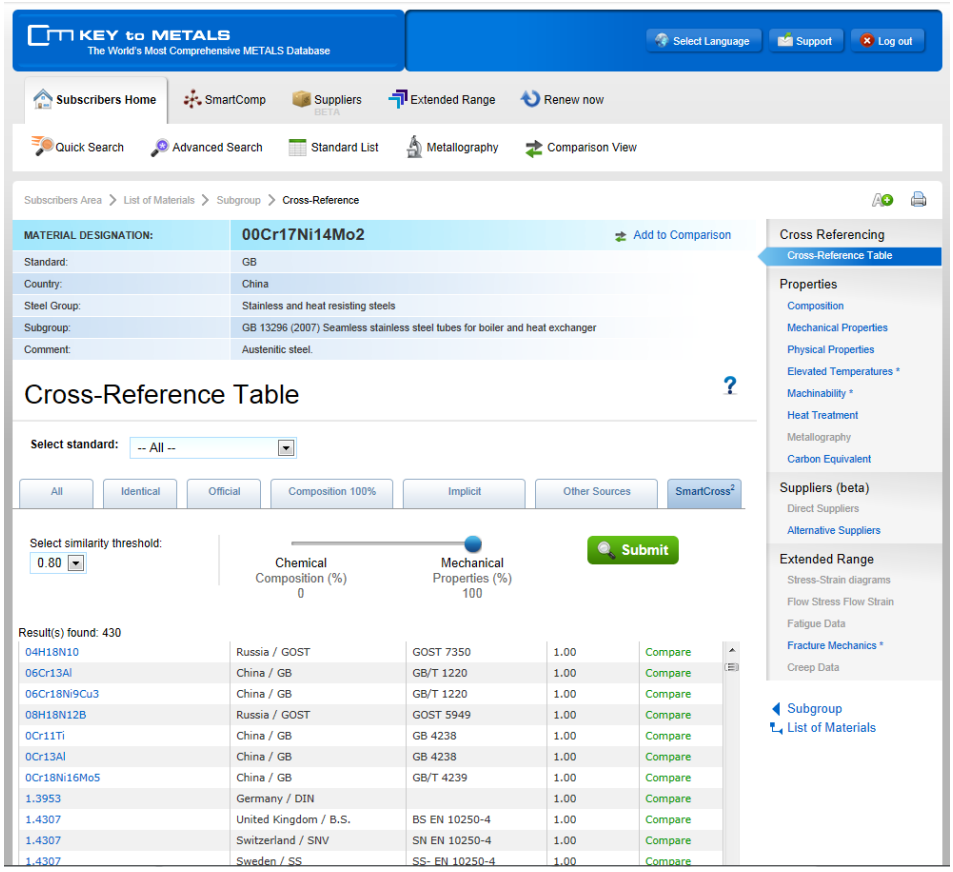

Fig. 5 An extension of the model to mechanical similarity of alloys and its combination with chemical similarity.

Mechanical similarity can provide better comparison results in many cases where the similarity of mechanical properties is more important than the similarity of compositions, in the field such as structural calculations and Finite Element (FE) analysis. After being implemented as a Web service within KEY to METALS, in the first 6 months only it has been used by many companies from the field of metal working, machinery, automotive, oil industry, utilities, and academia from over 40 countries, including some of the leading Computer Aided Engineering (CAE) companies and vendors worldwide.

Conclusions

In this model, a statistical approach has been applied onto a large alloy properties database to estimate the missing limits for alloying elements. The resulting "virtual content" is then used as in input for the fuzzy-logic based algorithm and again applied to large alloy properties database to achieve successful comparison and cross-referencing of metals. Comparing to the standard approach in finding equivalent materials which involves manual search by human experts or relies on the

recommendations of SDOs, the model proved to be able to provide better and more diversified results, in a considerably more practical and faster way, especially when combined with a user-friendly software interface.

Future research will be oriented to further refinements of the model by a controlled biasing of some of the alloying elements in the virtual content. This is a consequence of the fact that the content of some elements is indeed in reality biased for economical reasons. For example, the content of chromium in stainless steel tends to be closer to the lower limits due its high price, meaning producers try to minimize it in order to achieve substantial cost savings. The opposite is true when alloying elements are cheaper than the base element.

More challenging but a potentially very attractive direction of further research is including more mechanical and physical properties into the model, which can open an opportunity to make a simultaneous multi-parameter comparison of different properties and define ad-hoc relative importance of the properties depending on the concrete application. A preliminary research has indicated that this kind of model could be applicable to other structural materials beyond metals, such as polymers, ceramics and composites.

Also, this kind of combined similarity can be used for predicting "non-standard" advanced material properties, such as fatigue, stress-strain curves and fracture mechanics parameters, which are becoming essential for the most demanding modern structural calculations, FE analysis and simulations, by using known properties of similar materials. An initial research [14] has shown that the model and especially the method for calculating missing property values are fully applicable to the mechanical properties of metals. The calculations of some advanced properties such as fatigue based on this model have also shown to be at least comparable or better in comparison to the standard methods [15]. However, a considerable research effort is still needed in further development of the general model, possibly involving some new computational techniques such as neural networks for modeling properties [16, 17]; even more effort may be needed for validating the developed model, since the validation of model which includes a higher number of mechanical and advanced properties becomes significantly more complex.

References

[1] Q. Wang, A. Scott, Int. J. Adv. Manuf. Technol., 32 (2007), p 843–855
[2] F. Cverna, Worldwide Guide to Equivalent Irons and Steels, ASM International, Cleveland, 2006
[3] F. Cverna, Worldwide Guide to Equivalent Nonferrous Metals and Alloys, ASM International, Cleveland, 2007
[4] P. Cunat, http://www.euro-inox.org/pdf/map/AlloyingElements_EN.pdf, accessed 1 August 2012
[5] H. Chandler, Metallurgy for Non-metallurgist, ASM International, Cleveland, 1998
[6] V.Pocajt, Metalurgija-Journal of Metallurgy, 11 (2005), p 95-106
[7] S. Pocajt, V. Pocajt, K. Raić, In: Proceedings of First Metallurgical & Materials Engineering Congress of South-East Europe, 2013, p. 129-134.
[8] V. Pocajt, SmartMetals Method and System to Identify Metal Alloys, United States Patent No US 2009/0076739 A1

[9] V. Pocajt, Z. Ševarac, A. Kovačević, J. Chemometrics, 11 (2009), p 555-561
[10] Key to Metals AG, http://www.keytometals.com, accessed 8 June 2013
[11] Thermo Fisher Scientific, OXSAS Analytical Software User Manual, Thermo Fisher Scientific, Ecublens, 2009
[12] X. F. Zha, Int. J. Adv. Manuf. Technol., 25 (2005), p 233–243
[13] H. Zimmermann, Fuzzy Set Theory and Its Applications, Springer Verlag, 2001
[14] V. Pocajt, V. Ogarević, In: Proceedings of 33 Convegno Nazionale AIM, Brescia, 2010
[15] M. Roessle, A. Fatemi, Intern. J. Fatigue, 22 (2000), p 495–511
[16] K. V. Sudhakar, M. E. Haque, J. Mater. Eng. Perform., 10 (2001), p 31-36
[17] N. Koksal, Comp. Mat. Sci., 2009, 47, p 86–92