

## Use of computational intelligence for the prediction of vacancy migration energies in atomistic kinetic monte carlo simulations

**Nicolas Castin**\*

*Physique des Solides Irradiés et des Nanostructures (PSIN), Université Libre de Bruxelles (ULB),  
boulevard du Triomphe CP234, 1050 Brussels, Belgium.*

**Roberto Pinheiro Domingos**

*European Centre for Soft Computing, Calle Gonzalo Gutiérrez Quiróz s/n  
Mieres, Asturias 33600, Spain*

**Lorenzo Malerba**

*NMS, SCK•CEN, Boeretang 200  
B2400, Mol, Belgium  
E-mail: [lmalerba@sckcen.be](mailto:lmalerba@sckcen.be)  
[www.sckcen.be](http://www.sckcen.be)*

Received: 24-04-2008

Revised: 02-10-2008

### Abstract

In this work, we try to build a regression tool to partially replace the use of CPU-time consuming atomic-level procedures for the calculation of point-defect migration energies in Atomistic Kinetic Monte Carlo (AKMC) simulations, as functions of the Local Atomic Configuration (LAC). Two approaches are considered: the Cluster Expansion (CE) and the Artificial Neural Network (ANN). The first is found to be unpromising because of its high computational complexity. On the contrary, the second provides very encouraging results and is found to be very well behaved.

*Keywords:* Neural Networks, Fuzzy Logic, Cluster Expansion, Vacancy Migration Energy

### 1. Introduction

Phase transformation in general, and solute precipitation in particular, are spontaneous physical phenomena that may occur during operation in structural materials, e.g. steels, and dramatically modify their mechanical properties, thereby threatening the safety of the affected component. Models reliably describing the kinetics of these phenomena are therefore of importance for the safe exploitation of industrial nuclear power plants. For example, the formation of copper-rich precipitates and nanovoids under neutron irradiation is widely accepted

to be the main cause of hardening and embrittlement of nuclear Reactor Pressure Vessel (RPV) steels during operation<sup>1</sup>, as a consequence of their acting as obstacles to dislocation motion. Experimental evidences (see e.g. Refs. 2-4) have highlighted that any model for the prediction of RPV steel hardening versus radiation dose (which is the basic requirement for the RPV lifetime assessment) needs to be able to account as correctly as possible for the build-up of Cu precipitate and Cu-vacancy complex density.

---

\* Contact address : Studie Centrum voor KerneEnergie•Centre d'études de l'énergie nucléaire (SCK•CEN), Boeretang 200, B2400, Mol, Belgium.  
[ncastin@sckcen.be](mailto:ncastin@sckcen.be)

N. Castin et al.

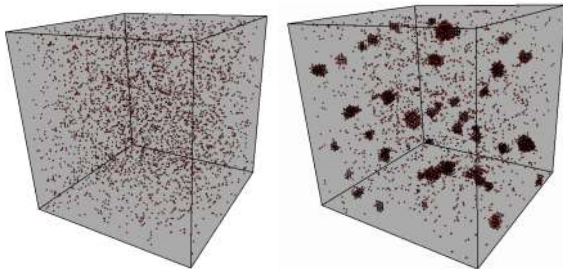


Fig. 1. AKMC simulation box with about 700,000 atoms (1.4% Cu). (left) Initial state with the copper atoms randomly distributed. (right) An intermediate state (after billions of AKMC steps).

Atomistic Kinetic Monte Carlo (AKMC) simulations<sup>5-11</sup> are among the best suited tools for studying the precipitation of Cu in Fe via a vacancy migration mechanism, as a subpart of the general study of RPV steels long-term evolution after decades of operation. AKMC is a compromise between Molecular Dynamics (MD), that considers events at the atomic time and length scale, and coarse-grained tools, such as Object KMC<sup>12</sup> and rate theory<sup>13-14</sup>, that are necessary to extend the simulation to the macroscopic scale. AKMC techniques retain the atomic level description, but reduce the number of possible events to the very basic mechanisms of single-defect diffusion and can thus encompass a timeframe (much) larger than MD.

Figure 1 shows an example of AKMC simulation. The cubic box is filled with matrix Fe atoms, and contains a small percentage of Cu. The atoms are arranged in a 3D rigid grid of coordinates that corresponds to the Body Centred Cubic (BCC) crystallographic structure. Several sites are however empty, corresponding to vacancies. At each step of the calculation, one of them is moved as shown on figure 2. Each vacancy has eight possible destinations, corresponding to the eight corners of the BCC cells. One of all candidate jumps, whose number is eight times the number of vacancies present in the system, is selected according to its probability, using the Monte Carlo sampling method<sup>15</sup>.

The vacancy jump probability,  $p_j$ , is calculated using the classical transition rate theory, i.e. using an Arrhenius-like expression for the jump frequency, that describes the jump as a thermally activated process:

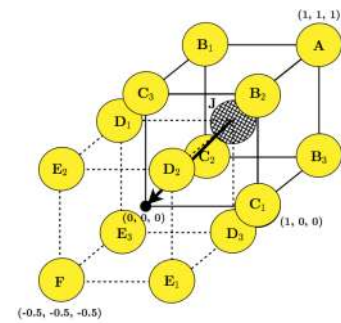


Fig. 2. One AKMC step corresponding to the migration of one vacancy (in reality, it is rather one of its neighbouring atoms that migrates to it). The figure shows the migration of the hatched atom to the vacancy, situated on the front bottom left corner of its BCC cubic cell, in plain lines. The dashed line shows the BCC cell of destination.

$$p_j = \frac{v_{0,j} \exp(-E_j/k_B T)}{\sum_k v_{0,k} \exp(-E_k/k_B T)} \quad (1)$$

Here  $v_{0,j}$  is a prefactor that is assumed to be constant and of the order of the Debye frequency for Fe (in the present case),  $k_B$  is Boltzmann's constant,  $T$  is the absolute temperature and  $E$  is the Vacancy Migration Energy (VME), which is the key parameter to be assessed and the focus of the present work. A precise definition and description of the latter are given in section 2. For the moment, it is important to know that the VME depends on the Local Atomic Configuration (LAC) and that it can be rigorously calculated with MD-type tools.<sup>†</sup> This is, however, a very time-consuming operation, that cannot possibly be undertaken at every AKMC step. Our project is thus aimed at partially replacing this rigorous calculation by a regression tool, trained to predict the VME on the basis of a limited number of MD calculated examples. Two possibilities have been envisaged and are reported in sections 3 and 4. The objective is to be able to calculate the VME hundreds of billions of times within a reasonable time

<sup>†</sup> Note that, technically speaking, it is not full MD that is used for the VME calculation, but algorithms based on the use of an interatomic potential that are easily implemented in an MD code. There exist a number of them (see Ref. 9), but here we shall not enter the detail of these algorithms and will generically speak, for simplicity, of MD calculations. Note also that calculations of the same type, more reliable although much more expensive in terms of CPU-time, can also be performed using *ab initio*, i.e. quantum-mechanics-based, methods.

frame, with as few simplifications as possible, because of the crucial role that the VME plays in the time increment of the AKMC, and thus on the prediction of the kinetics of the simulated process.

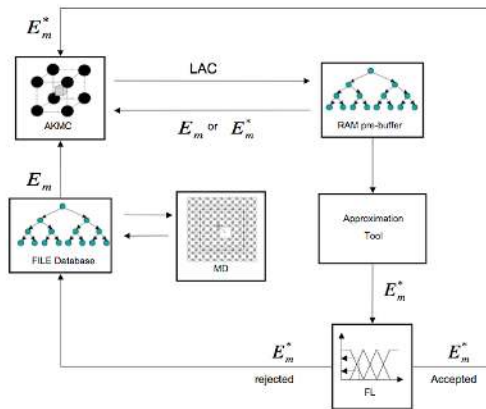


Fig. 3. The evolutionary VME prediction system.  $E_m$  is the MD calculated VME, whereas  $E_m^*$  is the prediction made by the approximation tool. The pre-buffer and the VME database are implemented under the form of a binary tree.

In order to keep the approximation tool qualities into control, a Fuzzy Logic (FL) based risk assessment method has been developed, too, to determine the probability that a VME prediction is subjected to an unreasonably high error. Such a tool allows the construction of the evolutionary system shown on figure 3. The rejected VME are MD calculated, before adding this new information in the existing database. Once the latter has been filled with a certain amount of new entries, both the approximation tool and its FL risk assessment module can be automatically re-trained. Such a strategy is promising to enhance the VME prediction qualities during the AKMC. The FL module is described in section 5.

Physical considerations suggesting the convenience of this scheme and a few preliminary results, obtained with first rudimentary algorithms, with only a few hints about the architecture of the numerical tools, have been already reported in Ref. 16. Here we focus on the

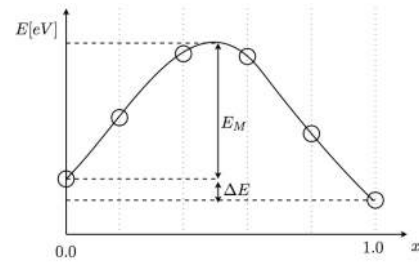


Fig. 4. Vacancy Migration Energy (VME).  $x$  is the dimensionless advancement coordinate along the atom displacement path. The y-axis of the figure is the total energy of the system in electron volts.  $E_M$  is a cubic spline interpolation of the saddle point in the minimum energy path found by the drag method.

detailed description and discussion of the algorithmic part, presented in its latest form.

## 2. The Vacancy Migration Energy

The VME is the difference between the largest energy encountered during the process leading to the exchange between a vacancy and a nearby atom (saddle point) and the initial energy of the system, as illustrated on figure 4. The VME can be estimated in a number of ways. Empirical formulas based on the total energy difference ( $\Delta E$ ) (see Ref. 5, 6, 9-11, 17.) are the simplest to apply, but also the most approximate ones, as discussed also in Ref. 16. Rigorous calculations can be undertaken with methods such as “drag”, “dimmer” or “Nudged Elastic Band” ones. The interested reader can find a general survey on that topic in Ref. 18. The method we used was a drag refined with cubic splines interpolation. The total energy of the system is calculated with MD performing a quench of the crystal.

The VME varies with the Local Atomic Configuration (LAC), as illustrated on figure 2. The A, B1, ... , F atoms shown in addition to the migrating one are the first nearest neighbors (1nn) of both its initial and final positions. They can be of several chemical types, or even be another vacancy. Depending on their nature, the corresponding VME will be different. The LAC can thus be coded under the form of an array of integers :

$$1nn \text{ LAC} = [ J \ A \ B1 \ B2 \ \dots \ E2 \ E3 \ F ] \quad (2)$$

N. Castin et al.

Each entry corresponds to one particular site around the jump and the value it takes depends on the nature of the

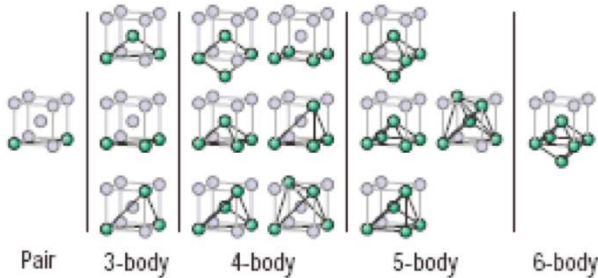


Fig. 5. Types of many-body interactions considered for the 3nn approximation.

object found therein. The length of the LAC string depends on the accuracy of the correlation. For example, the 1nn approximation shown on figure 2 considers 15 atomic sites. Taking further nearest neighbors into account makes this number increase. The list of possible values that the integers may take depends on the physical problem at hand. We talk about the FeAB\_Xnn problem when, in addition to the Fe matrix and 1 migrating vacancy, both the A and B species may be encountered and participate in defining the LAC and when the Xnn approximation is used, i.e. the LAC is extended to include the Xth shell of neighboring sites.

Table 1 summarizes the problems of interest for the study of the FeCu binary alloy. It is clear that the number of possible LACs explodes quickly and that a full calculation with MD is totally unfeasible. Note that, in practice, the inherent symmetries of the BCC configuration allow the total number of LACs to be reduced by a factor 6 at the most. This “trick”, however, obviously does not remove the underlying complexity mentioned earlier.

Our problem is consequently to correlate the VME with respect to 15 to 39 integer (categorical) type input variables. The output is a smooth real type function that takes values from 0.0<sup>‡</sup> to, say, 1.5 eV, with the largest accuracy possible. The parameters of the drag and

<sup>‡</sup> Negative values are meaningless as when introduced in (1) they may lead to unreasonably large jump probabilities. Physically, negative values would imply that the transition is not thermally activated, but spontaneous, which means that the initial state is ill-chosen.

Table 1. Problems of interest for our application. The bold values in the last column are those for which a full calculation with MD and tabulation is feasible. As an order of magnitude, one VME calculation with MD takes about one minute on modern personal computers.

Problem	Num sites	Num LACs
FeCu1nn ; FeVac1nn	15	$2^{15} = \mathbf{32768}$
FeCu2nn ; FeVac2nn	21	$2^{21} = \mathbf{2097152}$
FeCu3nn ; FeVac3nn	39	$2^{39} = 5.50e+11$
FeCuVac1nn	15	$3^{15} = \mathbf{1.43e+7}$
FeCuVac2nn	21	$3^{21} = 1.05e+10$
FeCuVac3nn	39	$3^{39} = 4.05e+18$

quenching methods have been selected as a compromise between speed and accuracy.

### 3. Cluster Expansion

A cluster expansion for an alloy can be viewed as a generalized Ising Model<sup>19</sup>. Any property of the alloy that only depends on the atomic configuration, the total energy in particular, may be expressed by means of such an expansion. Its application to energy barriers (VMEs) has been proposed e.g. in Ref. 20. More precisely, the occupation variables of the LAC allow a description of the energy barrier as an expansion in terms of polynomials :

$$E_{CE} = J_0 + \sum_{sites} J_i S_i + \sum_{pairs} J_{ij} S_i S_j + \sum_{triplets} J_{ijk} S_i S_j S_k + \dots \quad (3)$$

where  $S_i$  are the LAC entries and  $J_{ij}$  are the coefficients of the basis functions that can be fitted e.g. to the MD calculated VME's for a variety of different LAC's. Figure 5 shows the types of many-body interactions considered in an expansion, up to the 3nn. The  $J_{ij}$  coefficients thus represent the contribution of each of these interactions to the VME.

Of course, only a finite number of interactions can be involved in the expansion. The choice of which

interaction is more essential to a particular expansion is an open question.

The fit of the expansion of each LAC to the corresponding MD calculated VME is performed by adjusting the coefficient  $J_{ij}$ , minimizing the least-square error between the calculated energy and the predicted energy proposed by the expansion. This minimization can be undertaken with a traditional optimization method like Single Value Decomposition<sup>21</sup> (SVD) based on pseudo-inverse matrix, stochastic methods like Genetic Algorithms<sup>22</sup> (GA), or non-linear parameter fitting like Levenberg-Marquardt<sup>23</sup> (LM).

It was decided, as a first approach, to study the performance of the cluster expansion method for the simple binary 1nn problems. Even for this simple situation it was necessary to impose a truncation on the expansion by considering only many body interaction consisted of pairs or triplets of atoms, in order to avoid the number of terms to explode.

### 3.1 Binary problems

This section presents our results with the FeCu1nn and FeVac1nn problems (see Table 1), where the  $S_i$  occupation variables may thus take only two different values. The number of free parameters to be optimized was  $15 + C_{15}^2 + C_{15}^3 = 575$ . The percentage of points used for training was 20%. The optimization method applied in this work was GA (SVD and LM gave similar results).

Figure 6 shows the results obtained with the only consideration of pairs. The cluster expansion VME predictions are clearly well behaved for the FeCu1nn problem, even with such a simple model. The additional consideration of triplet interaction, however, allows to reduce the average error committed, as shown on figure 7. On the contrary, the cluster expansion predictions quality is much less satisfying for the FeVac1nn problem.

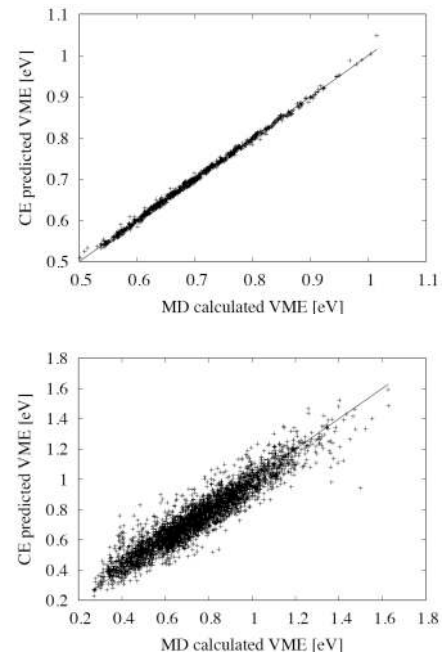


Fig. 6. Cluster expansion VME predictions for the FeCu1nn (up) and FeVac1nn problems (down), with only pair interactions taken into consideration. The average errors are respectively 0.53% and 7.12%.

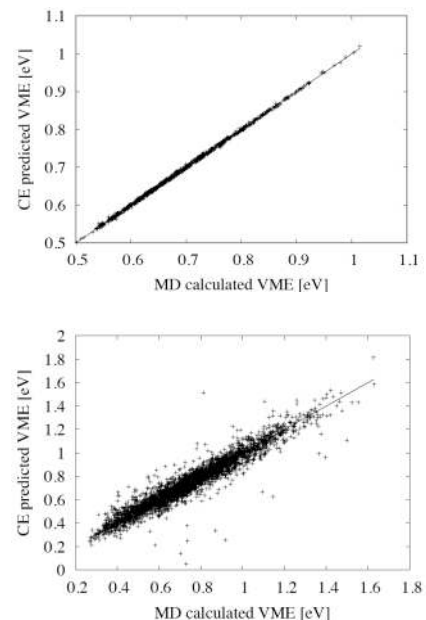


Fig. 7. Cluster expansion VME predictions for the FeCu1nn (up) and FeVac1nn problems (down), with pair and triplet interactions taken into consideration. The average errors are respectively 0.29% and 3.61%.

N. Castin et al.

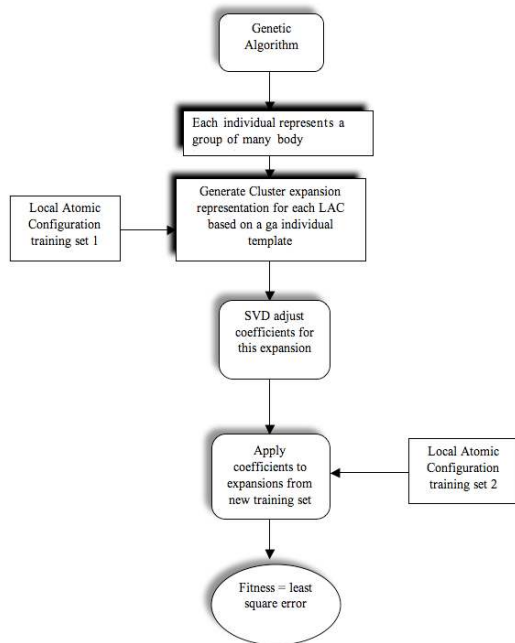


Fig. 8. GA based Cluster Expansion optimization.

### 3.2. Ternary problem

The problem considered here is the more general and more complicated FeCuVac1nn (see Table 1), where both copper atoms and vacancies may be encountered in the LAC.

A solution was imagined to find a way to take into account the possibility for some interactions, in the cluster expansion formulation, to be more important than others<sup>24</sup>. Of course there is no way to identify these interactions on-the-fly during the optimization process, so it was decided to design a GA based intelligent system, depicted on figure 8, in order to identify the relevant many-body interactions of a given problem and to obtain the adjustment of the coefficients for these interactions.

The individuals of the GA population represent different ways to construct a cluster expansion, or different many body interactions to consider. A population of different templates, or possible cluster expansions, is created in the first generation. A training set consisting of local atomic configurations is presented to each individual (training set 1), and these configurations are “translated”

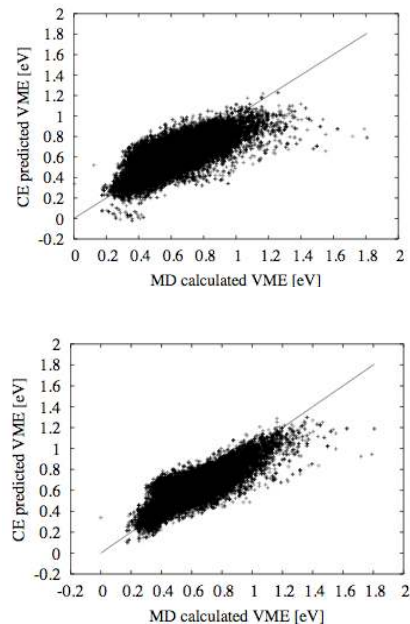


Fig. 9. Cluster expansion VME predictions for the FeCuVac1nn with pairs interactions (up) and with triplets interactions (down). The average errors are respectively 15.72% and 11.74%.

to a cluster expansion according to the scheme of this individual. A SVD algorithm finds the appropriate coefficients, minimizing the least square error on the training set 1. With the adjusted coefficients, a second training set (training set 2) is translated as the cluster expansion proposed by the individual, the least square error is calculated and used as a measure of how good this individual, or this particular set of many-body interactions, is able to produce a good prediction. As the genetic algorithm evolves, only individuals that represent suitable expansions would survive.

The results obtained with the FeCuVac1nn problem are shown on figure 9. The performance of the GA based model is surprisingly good compared to the preliminary results we obtained with the traditional model.

### 3.3. Conclusion on CE

The Cluster expansion approximation tool has been successfully applied to simple VME prediction problems. A GA-based model has been devised to



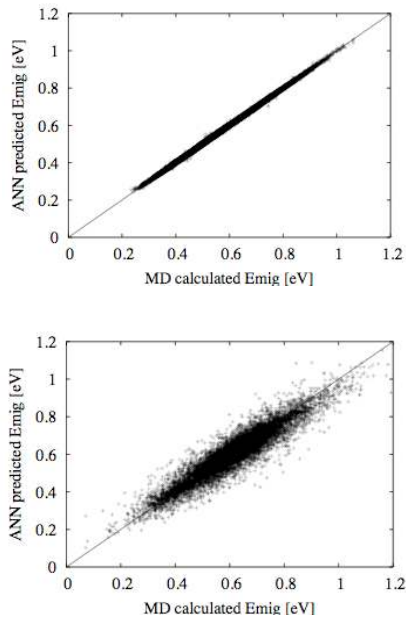


Fig. 10. ANN Prediction qualities for the FeCu3nn (left) and for the FeCuVac3nn (right) (maximum 7 vacancies) problems.

determine the most important interactions to be taken into account, enabling the number of terms in the expansion to be limited.

However, the potential of the cluster expansion seems limited because of the high complexity that the optimization problem takes when the number of atomic sites to be considered is increased. For example, the consideration of the 2nn problems (21 sites in the LAC) requires the determination of around 10,000 free parameters if triplet interactions are introduced in the expansion. An optimization problem of this size is known in GA terminology as Large Parameter Optimization Problem (LPOP) and requires a large population size and many generations to converge. The computational complexity, therefore, quickly explodes. For this reason and considering that, at the same time, we were obtaining better results with a hybrid fuzzy – neural network framework (see following sections), it was decided to abandon this model in favor of a more efficient and robust approach.

#### 4. Artificial Neural Networks

Artificial Intelligence (AI) is the combination of algorithms, data and software used to develop computer systems that can be said to be *intelligent*. Here, the defining feature of intelligence is the capability of learning from past experience and solving problems when important information is missing, so as to be able to handle complex situations and to react correctly to new ones. There are many different computational models that are considered branches of the artificial intelligence field, each one suitable to a different kind of problem. For our particular application, the feed-forwards Artificial Neural Network (ANN) is particularly well suited, as it provides a general framework for representing non-linear functional mappings between a set of input variables and a set of output functions<sup>25</sup>. It is a *universal approximator* in the sense that a Multi-Layer Perceptron (MLP) can approximate any continuous multivariate function to any desired degree of accuracy, provided that a sufficiently large number of hidden neurons are available<sup>26-27</sup>.

##### 4.1. Predictions quality

Figure 10 shows the ANN prediction qualities for binary and ternary 3nn problems. The average error committed is 0.51% for FeCu3nn and 3.37% for FeCuVac3nn (with maximum 7 vacancies in the LAC). The correlation coefficient  $r^2$  is larger than 0.99 in both cases. The ANN is thus clearly outperforming the cluster expansion, not only because the error committed is much smaller, but also because the training procedure is much less computational time demanding.

The next sections present the experiments we made to study different ANN architectures and training algorithms.

##### 4.2. Experimental conditions

Two MLP architectures have been considered. The first is the classical fully interconnected mono-hidden layer network without bypass connections, using sigmoid

N. Castin et al.

activation functions, as widely described in Ref. 25-26 and by many others. It will be denoted as Fixed Architecture MLP (FAMLP) from now on. The second is Fahlman-Lebriere's Cascade Correlation Network (CCN) as described in Ref. 28. It is, contrary to the FAMLP, a constructive algorithm where hidden units are added in successive layers.

Two algorithms have been considered for the FAMLP training. The first one is the steepest-descent Resilient Propagation (RPROP) used in batch mode, as described in Ref. 29. The second one is the Levenberg-Marquardt (LM), described for example in Ref. 25, which is an approximation of the second order Newton method and that does not require the computation of the Hessian matrix. The synaptic connections were initialized at random between  $\pm 2.4/F$  ( $F$  is the node fan-in<sup>26</sup> as recommended in Ref. 26).

The CCN training algorithm was changed compared to the original Fahlman-Lebriere's. Instead of proceeding to the addition of a new hidden node in two phases, all synapses linked to it are trained all together with the output ones, with a classical training algorithm (chosen to be LM). The reason is that it seems to us that the original Fahlman-Lebriere training scheme is best suited for classification problems using the 1-of-c coding for the output signal. Furthermore, the QuickProp<sup>30</sup> algorithm originally proposed for CCN training<sup>28</sup> didn't give more satisfactory results than RPROP and LM and is therefore not considered in the present paper. 20 candidate nodes were considered before any new hidden unit addition. The synapses initialization strategy was the same as for FAMLP training, and the activation functions were chosen at random amongst the sigmoid, Gaussian and hyperbolic tangent.

### 4.3. Experiment 1

All architectures and training algorithms have been tested on the FeCu1nn and FeCu2nn problems. The FAMLP was trained with several numbers of hidden nodes. The experiments were run 20 times.

Figure 11 shows the experiments for the FeCu2nn problem. Figure 12 summarizes the results for both the

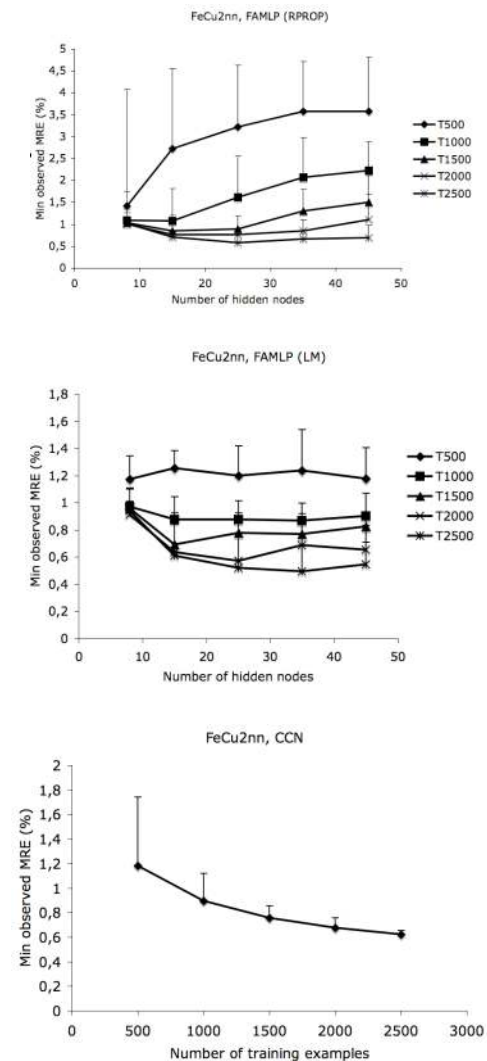


Fig. 11. FeCu2nn trained with FAMLP and CCN. The error bars show the max observed MRE over the 20 trainings performed. T is the number of training examples.

FeCu1nn and the FeCu2nn problems. The following observations can be made :

- The best-suited number of hidden nodes for the FAMLP is not easy to determine, because of the substantial variance of the final MRE. In fact, it is absolutely necessary to run all experiments several times, which is bad news from the computational point of view.



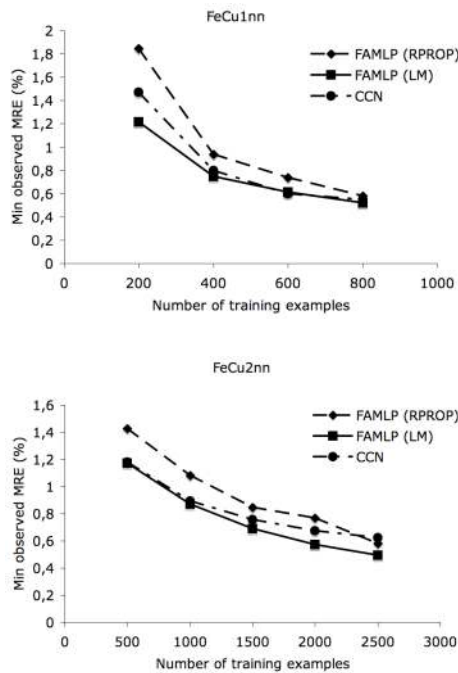


Fig. 12. Minimum MRE observed amongst all training experiments performed.

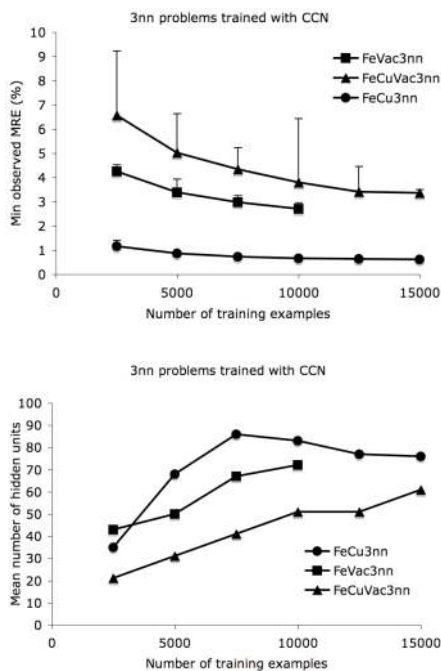


Fig. 13. Experiments with the 3nn problems.

- For FAMLP, the RPROP algorithm does not reach as low MRE as LM, except if the training set is large enough. LM is thus always preferable<sup>§</sup>.
- The final MRE variance is unsurprisingly lower for CCN than FAMLP.
- The CCN architecture is in general not capable of reaching the same MRE as the FAMLP, despite its advantage to construct the network automatically. In-depth architectures are thus apparently less appropriate to the VME correlation problem than the use of one single hidden layer.

It is thus clear that reaching a reasonably low MRE is rather easy, whereas fine-tuning is not conceivable without quite a number of training experiments.

#### 4.4. Experiment 2

Figure 13 shows the experiments performed with the 3nn problems. Only the CNN architecture was tested. The experiments were run 20 times.

The comparison of figure 12 and 13 shows clearly that the ANN need in training examples to converge to the lowest possible MRE behaves well with the problem complexity. Shifting from 1nn to 2nn or 3nn has not made this number explode. The same observation holds for the comparison of the FeVac and FeCuVac problems to the FeCu one. This is a very important point vis-à-vis the extension of the methodology to more complex problems.

#### 4.5. Conclusion on ANN

The ANN is clearly a very promising tool for the VME regression versus LAC. Low mean residual errors of predictions and very good correlation coefficients are indeed very easily obtained. However, ANN fine-tuning is not an easy issue, because quite a lot of training

<sup>§</sup> In practice, LM requires much less training epochs to converge than RPROP or other first order methods. However, The actual training time can be reversed for large systems because of the complexity of the LM :  $\mathcal{O}(W^3+TW^2)$  where W is the number of synapses and T the number of training examples. Anyway, LM was by far the fastest for our application.

N. Castin et al.

experiments are in practice necessary before the best possible ANN performance can be reached.

Consequently, it seems to us that the CCN training scheme is a good starting point to make the first NN training trials on a new problem, in order to determine the number of training points required to reach the lowest possible MRE. Then a long series of FAMPLP trainings, with the LM algorithm applied on different network architectures, are to be performed as a second step for fine-tuning.

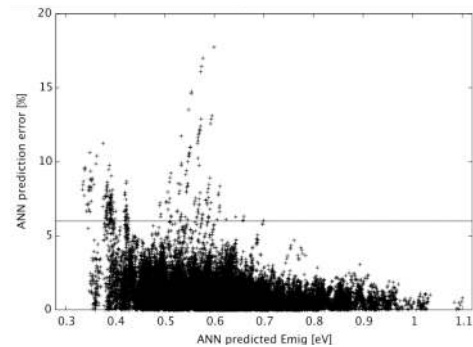


Fig. 14. ANN trained for the FeCu1nn problem. The MRE is 1.06%. The ERT is set at 6%.

### 5. Risk assessment on the VME approximation

As already mentioned, a Fuzzy Logic<sup>31</sup> (FL) system has been designed to assess the uncertainty<sup>32</sup> inherent to the use of the ANN and to evaluate the "risk"<sup>33</sup> associated with its use instead of the full calculation, so as to be able to build an integrated system, capable of feedback. Figure 14 shows an example of ANN trained for the FeCu1nn problem and of associated risk assessment scheme. Despite the reasonably low MRE, the error committed for some cases may be rather large, up to 17.5% for this particular example. The objective of the FL would then be to identify, on the basis of both the LAC and the corresponding predicted VME, whether the correlation error is probably larger than a certain Error Rejection Threshold (ERT).

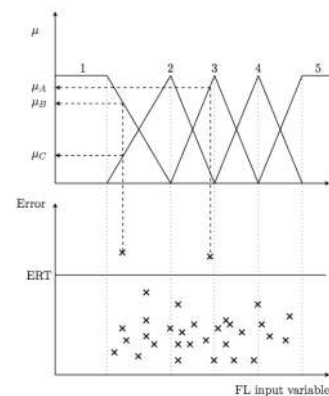


Fig. 15. FL sets and rules determination by the GA.  $\mu$  is the relevancy to belong to a certain FL set.

The FL system we developed follows a Sugeno model<sup>34-35</sup> and produces an output that is either 0 or 1, respectively meaning "acceptance" or "rejection" of the ANN predicted VME. The FL inputs are various information extracted from the LAC. For example, the variables for the FeCu1nn problem were :

- The total number  $N_{Cu}$  of copper atoms in the LAC.
- The difference  $\Delta C_{Cu}$  between the number of Cu atoms that are 1nn of the jumping atom and the number of Cu atoms that are 1nn of the jumping vacancy.
- The ANN prediction  $E_m^*$  of the migration energy.

Triangular-shaped sets are used for the fuzzification part of the Sugeno model.

An automated Genetic Algorithm (GA) based learning scheme has been developed in order to generate the FL system definition on the basis of the full VME available database. The GA optimization variables are the central coordinates of the triangular sets on the FL inputs, as shown on figure 15. The rules' conclusions are determined after a passage in the VME database. If at least one point that has an unacceptable error fulfils a certain rule, its conclusion is automatically chosen to be 1 ("rejection"). The conclusion is 0 ("acceptance") otherwise. The GA objective function is :

$$f = (1 + R_1) \cdot (1 + R_2) + 1 \quad (4)$$

Where  $R_1$  is the proportion of acceptable ANN predictions wrongly rejected by the FL and  $R_2$  is the proportion of unacceptable predictions wrongly

accepted. The GA task is thus to select the fuzzy sets as properly as possible, so as to isolate in what conditions precisely is the ANN failing at producing a good VME prediction.

The classical Sugeno inference scheme has been modified in order to improve the FL performance. The FL output  $O$  is calculated by :

$$O = \text{round}(o); o = \frac{\sum_r H_r \cdot C_r \cdot T_r}{\sum_r H_r \cdot T_r} \quad (5)$$

Where  $H_r$  is the rule relevancy and  $C_r$  is its conclusion. The  $T_r$  coefficient (named trust index) has been added. It is the maximum relevancy observed by the GA that has led to the conclusion  $C_r$  for the rule  $r$ . For example, on figure 15, the conclusion  $C_r$  of all rules involving the set number 3 on the FL input considered is "rejection". Only one observation leads to that conclusion, but it was lying very close to the set peak. In that case, the "rejection" conclusion for the rule  $r$  can be used with a high degree of trust, and so  $T_r = \mu_a$  is very close to 1. On the other hand, the conclusion of all rules  $r^*$  involving the set number 2 are also "rejection", but with a much lower level of confidence. In this case,  $T_{r^*} = \mu_C$  is close to 0. The  $r^*$  rule has, consequently, fewer chances to induce wrong rejections of acceptable VME predictions.

The complexity of the GA objective function is rather large. Two passages in the overall VME database are indeed required : the first serves to determine the rules' conclusions (as depicted on figure 15) and the second serves to calculate the function  $f$ . The application of such a method is very time-consuming if the database is large, which is the case in practice for the 2nn and 3nn problems. In order to reduce that complexity, we have introduced a clustering operation on the FL learning data, depicted on figure 16. Points having a similar value for all FL inputs are removed, as long as they lie on the same side of the ERT. Only one point is left, and is assigned with a weight factor that corresponds to the

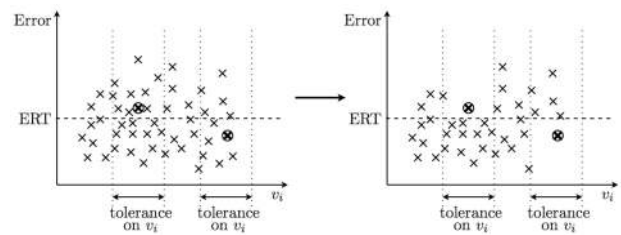


Fig. 16. Clustering operation on the FL learning data, around the two circled points, with respect to 1 FL input. In practice, points are removed only if the illustrated condition is fulfilled for all the FL inputs.

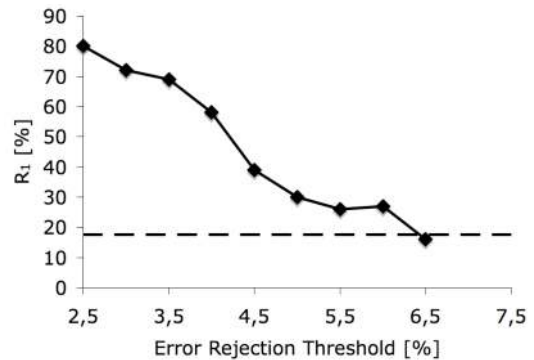


Figure 17 : Evolution with the ERT of the best-obtained  $R_1$  for the problem of Fe-Cu 1nn.  $R_2$  has immediately vanished during the GA optimization in all cases.

number of points removed plus one. This factor is then taken into account when the  $R_1$  and  $R_2$  members of the  $f$  function are calculated. In practice, points above the ERT are not clustered (it is affordable since they are not very numerous) and the clustering operation is very fast thanks to the help of a binary tree. The size of the FL learning table is in turn tremendously reduced and the GA optimization is consequently much faster. An appropriate choice of the selection tolerances shown on figure 16 allows the effect of clustering to be made negligible on the real FL abilities to isolate the unacceptable ANN predictions.

N. Castin et al.

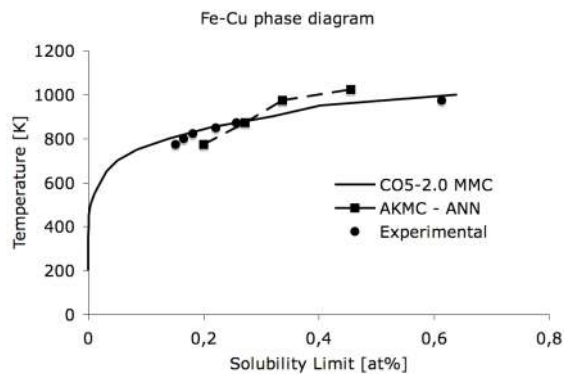


Fig. 18. AKMC results obtained with the use of the FeCu3nn trained MLP. The CO5 curve shows the Cu solubility limit of the inter-atomic potential used for the energy calculations.

The last undefined feature of the FL system is the appropriate selection of the ERT. In fact, we can see it as a compromise between ANN accuracy and AKMC speed. The ERT is actually not *a priori* chosen, but rather determined as low as possible with the constrain to keep  $R_l$  reasonably limited, as shown on figure 17. The choice of a large ERT is consequently less rigorous in the ANN risk assessment but summons MD calculations less often, and reversely.

## 6. Example of AKMC results with the use of ANN

Figure 18 shows the results of thermal annealing experiments computed with a small AKMC box containing 1.4% of Cu and 1 single vacancy. The VME was correlated with an MLP having a MRE of 0.5%. No FL Risk assessment was applied. Satisfactory predictions of the Cu solubility limit in Fe were obtained at different temperatures.

## 7. Conclusion and outlook

In this work, we have reported about our effort to develop a regression tool to partially replace a costly "molecular-dynamics" calculation of the local-atomic-configuration-dependent vacancy migration energy in an atomistic kinetic Monte Carlo scheme, where the local atomic configuration is presented under the form of an array of tens of categorical integers. We have in the first place envisaged a cluster expansion approach.

The latter has however been abandoned for an artificial neural network, that has been proven to be more robust, well behaved and promising for future developments of the project.

Our future objective is to continue the application of this method to more complicated situations. First, the number of atomic sites taken into consideration must be increased for a better description of the physics, and more chemical species are to be included in the model. Secondly, in order to tackle irradiation problems, the atomistic kinetic Monte Carlo model must be able to consider the migration of another type of point-defect, i.e. the self-interstitial. It is a more complicated event than the vacancy migration, mainly because of the more extended and anisotropic strain-field than for a vacancy, which enhances and complicates its interaction with neighboring atoms.

Our future work will thus have to face not only a larger number of artificial-neural-network input variables, but also a more complicated mapping between these inputs and the point defects migration energies that have to be predicted.

## Acknowledgments

This work was performed in the framework of the FP6 PERFECT project, partially funded by the European Commission under contract FI6O-CT-2003-508840. The authors wish to acknowledge F. Djurabekova and G. Cerchiara, who valuably contributed to the present work in its early stage.

## References

1. G.R. Odette and G.E. Lucas, JOM 53 (7) (2001) 8-22. Mater. Sci. Eng. 6 (1998) 19-28.
2. J.T. Buswell, W.J. Phythian, R.J. McElroy, S. Dumbill, P.H.N. Ray, J. Mace and R.N. Sinclair, J. Nucl. Mater. 225 (1995) 196-214.
3. K. Fukuya, K. Ohno, H. Nakata, S. Dumbill and J.M. Hyde, J. Nucl. Mater. 312 (2003) 163-173.
4. Y. Nagai, Z. Tang, M. Hasegawa, T. Kanai and M. Saneyasu, Phys. Rev. B 63 (2001) 134110.
5. B.D. Wirth and G.R. Odette, "Kinetic lattice Monte Carlo simulations of cascade aging in iron and dilute iron-copper alloys", Mat. Res. Soc. Symp. Proc. 540 (1999) 637-642.

6. C. Domain, C.S. Becquart and J.-C. van Duysen, *Microstructural Processes in Irradiated Materials*, Materials Research Society Symposium Proceedings, Vol. 650, edited by G.E. Lucas, L.L. Snead, M.A. Kirk Jr. et al. (Materials Research Society, Warrendale, PA, 2000), p. R3.25.
7. S. Schmauder and P. Binkele, *Comp. Mater. Sci.* 25 (2002) 174.
8. Y. LeBouar and F. Soisson, "Kinetic pathway from EAM potentials: influence of the activation barriers" *Phys. Rev. B* 65 (2002) 094103.
9. E. Vincent, C.S. Becquart and C. Domain, *J. Nucl. Mater.* 351 (2006) 88.
10. E. Vincent, C.S. Becquart and C. Domain, *Nucl. Instr. & Meth. in Phys. Res. B* 255 (2007) 78.
11. E. Vincent, C.S. Becquart, C. Pareige, P. Pareige and C. Domain, *J. Nucl. Mater.* 373 (2008) 387.
12. C. Domain, C.S. Becquart and L. Malerba, *J. Nucl. Mater.* 335 (2004) 121.
13. C. Ortiz, M.-J. Caturla, C.-C. Fu and F. Willaime, *Phys. Rev. B* 75, 100102(R) (2007).
14. C. Ortiz and M.-J. Caturla, *Phys. Rev. B* 75, 184101 (2007).
15. K. Binder (Ed.), "Monte Carlo Methods in Statistical Physics", Springer-Verlag, Berlin (1979).
16. F. Djurabekova, R. Domingos, G. Cerchiara, N. Castin, E. Vincent and L. Malerba, *Nucl. Instr. & Meth. In Phys. Res. B*, 255 (2007) 8.
17. Young66: W.M. Young and E.W. Elcock, *Proc. Phys. Soc.* 89 (1966) 735; Kang89: H.C. Kang and W.H. Weinberg, *J. Chem. Phys.* 90(5), 2824 (1989).
18. G. Henkelman, G. Jóhannesson, and H. Jónsson, "Methods for Finding Saddle Points and Minimum Energy Paths", in *Progress on Theoretical Chemistry and Physics*, 269-300, Ed. S. D. Schwartz (Kluwer Academic Publishers, 2000).
19. J.M. Sanchez, F. Ducastelle and D. Gratias, *Pysica* 128A (1984) 334-350.
20. Van der Ven, A; Ceder, G, "Vacancies In Ordered And Disordered Binary Alloy Treated With The Cluster Expansion," *Physical Review B*, Vol. 71, Issue 5, (2005).
21. Trefethen, L. N., & Bau, D. (1997). *Numerical linear algebra*. Philadelphia: Society for Industrial and Applied Mathematics
22. D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and machine learning*, Addison-Wesley, Reading, Mass., 1989.
23. Gill, P. E. and Murray, W. "Algorithms for the solution of the nonlinear least-squares problem", *SIAM J. Numer. Anal.* 15 [5] 977-992, 1978.
24. G. L. W. Hart, Volker Blum, Michael J. Walorski, and A. Zunger, "Evolutionary Approach for Determination of First-Principles Hamiltonians," *Nature Materials* 4, 391-394 (01 May 2005).
25. C.M. Bishop, "Neural Networks for pattern recognition", Clarendon press, Oxford, 1995.
26. S. Haykin, "Neural Networks: A Comprehensive Foundation", New York: MacMillian (1994).
27. K. Hornik, M. Stinchcombe and H. White, "Multilayer feedforward networks are universal approximators", *Neural Networks*, 2 (1989) 359-366.
28. Scott E. Fahlman and C. Lebiere, "The Cascade-Correlation Learning Architecture", in *Advances in Neural Information Processing Systems*, volume 2, pages 524-532, 1990.
29. M. Riedmiller and H. Braun, "A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm", *Proc. of the IEEE Intl. Conf. on Neural Networks*, pages 586-591, 1993.
30. Scott E. Fahlman, "An Empirical Study of Learning Speed in Back-Propagation Networks", *Computer Science Technical Report*, Carnegie-Mellon University, 1988.
31. L.A. Zadeh, *Fuzzy Sets & Systems* 100 (1999) 9.
32. D. Dubois and H. Prade, "Unfair Coins and Necessity Measures: Towards a Possibilistic interpretation of Histograms", *Fuzzy Sets & Systems* 10 (1983) 15.