Parametric Characterization of Hardness Profiles of Steels with Neuro-Wavelet Networks

V. Colla[1], L.M. Reyneri[2], and M. Sgarbi[3]

[1] Scuola Superiore Sant'Anna, Pisa (I), e.mail: vale@mousebuster.sssup.it
[2] Politecnico di Torino (I), e.mail reyneri@polito.it
[3] Scuola Superiore Sant'Anna, Pisa (I), e.mail: mirkosg@tin.it

Abstract. This work address the problem of extracting the Jominy hardness profiles of steels directly from the chemical composition. Wavelet and Neural networks provide very interesting results, especially when compared with classical methods. A hierarchical architecture is proposed, with a first network used as a parametric modeler of the Jominy profile, and a second one estimating parameters from the steel chemical composition. Suitable data preprocessing helps to reduce network size.

1 Introduction

Hardenability is a basic feature of steels: in order to characterize it, manufacturers usually perform the so-called *Jominy end-quench test* [1], which consists in measuring the hardness along a specimen of a heat-treated steel, at predefined positions; the measured values form the *Jominy hardness profile.*

Hardenability depends on chemical composition in a partially unknown fashion, therefore *black-box* models have been developed to predict the shape of Jominy profiles directly from chemical analysis. Most of them are linear, but this affects accuracy, especially when a wide variety of steels is considered.

Neural Networks (NNs) seem to cope well with such a modeling problem, as they are good approximators for strongly non-linear functions. An attempt to apply NNs to predict Jominy profiles has been made in [2] by using a standard Multi-Layer Perceptron (MLP) with one hidden layer, but there is no reported attempt to use Wavelet Networks, (WNs) for the same task.

Unfortunately, most methods based on NNs alone suffer from several caveats. For instance, their initialization and training requires a large amount of data, which are seldom easily and rapidly available. In addition, simple NNs may often predict profiles which are not physically plausible, unless very complex networks are used and long training processes are employed. It is therefore mandatory to accurately select the network structure, in order to obtain good performance, to reduce as much as possible the number of free parameters, and consequently to reduce the required size of the training set.

Another drawback of NNs alone is that no information related to physical characteristics of the steel can be extracted from the trained network; this means that NNs can only be used to predict the profiles themselves, but not any other steel characteristic.

This paper presents some more powerful methods based on two combined *Neuro-Wavelet Networks* (NWNs), where one network provides a parametric model of the Jominy profile, while the second one predicts the parameters as a function of chemical composition. The extracted parameters do have a strong relationship with the Jominy profile, of which they are a compact representation.

2 Neuro-Wavelet Unification

Radial Wavelet Networks are based on Wavelet decomposition and use radial Mother Wavelets $\Psi(\|\boldsymbol{X}\|) \in L_2(\Re^N)$ suitably *dilated* and *translated*. Such networks are based on *Radial Wavelons* (**WAVs**) which have a model based on the Euclidean distance between the input vector $\boldsymbol{X}$ and a *translation vector* $\boldsymbol{E}$, where each distance component is weighted by a component of a *dilation vector* $\boldsymbol{T}$:

$$y = \Psi\left(\sqrt{\sum_i \left(\frac{X_i - E_i}{T_i}\right)^2}\right) \tag{1}$$

A function $\Psi(\cdot)$ is admissible as a radial Wavelet only if its Fourier transform satisfies a few constraints not discussed here [5]. A commonly used function is the *Mexican hat* $\Psi(z) = (1 - 2z^2) \cdot \mathrm{e}^{-z^2}$.

Radial Wavelet Networks, as well as many other neural and fuzzy paradigms, can be viewed into a unified perspective by means of the *Weighted Radial Basis Functions* (**WRBF**) [4].

Each layer (array) of WRBF neurons is associated with a set of parameters: an *order* $n \in \Re$, defining the neuron's *metric* (mostly $n \in \{0, 1, 2\}$), a *weight matrix* $\underline{\underline{W}}$, a *center matrix* $\underline{\underline{C}}$, a *bias vector* $\boldsymbol{\Theta}$ and an *activation function* $F(z)$. The mathematical model of a WRBF neuron of order n (or, WRBF-n) is:

$$Y_i = F\left(\sum_j \mathcal{D}_n(X_j - C_{ji}) \cdot W_{ji} + \Theta_i\right) \tag{2}$$

where $F(z)$ can be any function (although in most cases monotonic functions or Wavelets or linear or polynomial functions are used) and the *distance function* $\mathcal{D}_n(\cdot)$ is defined as:

$$\mathcal{D}_n(X_j - C_{ji}) \triangleq \begin{cases} (X_j - C_{ji}) & \text{for } n = 0 \\ |X_j - C_{ji}|^n & \text{for } n \neq 0 \end{cases} \tag{3}$$

All the NWN paradigms used here have been re-conduced to WRBF in order to have common paradigms, methodologies, initialization strategy and learning rule, which is the main advantage of unification; *Radial Wavelons* are WRBF-2 neurons with $W_{ji} = (1/T_{ji})^2$, and $\underline{\underline{C}} = \underline{\underline{E}}$ (i.e. the matrix made of one translation vector $\boldsymbol{E}$ per neuron), while the activation function comes from the radial Mother Wavelet $F(z) = \Psi(\sqrt{z})$. Details on the unification of other neural paradigms can be found in [7, 4].

As far as initialization of NWNs is concerned, in this work we have used three forms of initialization:

- *Fixed initialization*: all weights, biases and centers are initialized to a predefined value (or set of values). This has been used for all the networks of the parametric model described in section 4.
- *Random initialization*: the parameters are initialized to random values (uniform distribution). This has been used for the WRBF-0 networks of the parameter estimator described in section 5.