Many processes show a non-linear static and dynamic behavior, especially if wide areas of operation are considered. Therefore, the identification of non-linear processes is of increasing interest. Examples are vehicles, aircraft, combustion engines, and thermal plants. In the following, models of such non-linear systems will be derived based on artificial neural networks, that had first been introduced as universal approximators of non-linear static functions.

20.1 Artificial Neural Networks for Identification

For a general identification approach, methods of interest are those that do not require specific knowledge of the process structure and hence are widely applicable. Artificial neural networks fulfill these requirements. They are composed of mathematically formulated neurons. At first, these neurons were used to describe the behavior of biological neurons (McCulloch and Pitts, 1943). The interconnection of neurons in networks allowed the description of relationships between input and output signals (Rosenblatt, 1958; Widrow and Hoff, 1960). In the sequel of this chapter, artificial neural networks (ANNs) are considered that map input signals $u$ to output signals $y$, Fig. 20.1. Usually, the adaptable parameters of neural networks are unknown. As a result, they have to be adapted by processing measured signals $u$ and $y$ (Hecht-Nielsen, 1990; Haykin, 2009). This is termed “training” or “learning”.

One can discern two steps in the design of a neural net. The first is the training, where the weights or other parameters of the neural net are optimized. The second step then is the generalization, where the net is used to simulate new data, that have not been part of the training data and allow to judge the performance of the net for unknown data. The goal is to obtain the smallest possible error for both training and generalization. The model error can be split in two parts as

$$E\frac{(y_0 - \hat{y})^2}{NUL} = E\frac{(y_0 - E\{\hat{y}\})^2}{TAB} + E\frac{E\{(\hat{y} - E\{\hat{y}\})^2\}}{OFO}.$$  \hspace{1cm} (20.1.1)
The bias error is a systematic deviation between the true system output and the expected model output, that appears when the model does not have enough flexibility to fit the real process (Underfitting). Consequently, the bias error decreases as the model complexity increases. The variance error is the deviation between the model output and the expected model output. The variance error increases as the number of degrees of freedom of the model increases. The model is more and more adapted to the specific peculiarities of the training data set such as noise and outliers. Hence, in choosing the model structure, there is always a trade-off between the bias error and the variance error, which is termed bias-variance dilemma, see Fig. 20.2 (German et al, 1992; Harris et al, 2002).

In identification, one is interested in approximating the static or dynamic behavior of processes by means of (non)-linear functions. On the contrary, if inputs and outputs are gathered into groups or clusters, a classification task in connection with e.g. pattern recognition is given (Bishop, 1995). In the following, the problem of non-linear system identification is considered (supervised learning). Thereby, the capability of ANNs to approximate non-linear relationships to any desired degree of accuracy is utilized. Firstly, ANNs for describing static behavior (Hafner et al, 1992; Preuß and Tresp, 1994), will be investigated, which will then be extended to dynamic behavior (Ayoubi, 1996; Nelles et al, 1997; Isermann et al, 1997).

**20.1.1 Artificial Neural Networks for Static Systems**

Neural networks are universal approximators for static non-linearities and are consequently an alternative to polynomial approaches. Their advantages are the requirement of only little a priori knowledge about the process structure and the uniform treatment of single-input and multi-input processes. In the following, it is assumed