Expert Systems With Applications 44 (2016) 265–274

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Numerical weather prediction revisions using the locally trained differential polynomial network

Ladislav Zjavka*

VŠB-Technical University of Ostrava, Faculty of Electrical Engineering and Computer Science, Department of Computer Science, Czech Republic

ARTICLE INFO

Keywords:
Polynomial neural network
Differential equation composition
Relative substitution derivative term
Multi-variable function approximation

ABSTRACT

Meso-scale forecasts result from global numerical weather prediction models, which are based upon the differential equations for atmospheric dynamics that do not perfectly determine weather conditions near the ground. Statistical corrections can combine complex numerical models, based on the physics of the atmosphere to forecast the large-scale weather patterns, and regression in post-processing to clarify surface weather details according to local observations and climatological conditions. Neural networks trained with local relevant weather observations of fluctuant data relations in current conditions, entered by numerical model outcomes of the same data types, may revise its one target short-term prognosis (e.g. relative humidity or temperature) to stand for these methods. Polynomial neural networks can compose general partial differential equations, which allow model more complicated real system functions from discrete time-series observations than using standard soft-computing methods. This new neural network technique generates convergent series of substitution relative derivative terms, which combination sum can define and solve an unknown general partial differential equation, able to describe dynamic processes of the weather system in a local area, analogous to the differential equation systems of numerical models. The trained network model revises hourly-series of numerical prognosis of one target variable in sequence, applying the general differential equation solution of the correction multi-variable function to corresponding output variables of the 24-hour numerical forecast. The experimental results proved this polynomial network type can successfully revise some numerical weather prognoses after this manner.

1. Introduction

Meso-scale meteorological models need as a rule to be forced by robust global numerical forecasts, which involve a large number of 3D matrix variables in several atmospheric layers, provided by lateral boundary conditions. Post-processing methods using local measurements can improve some numerical weather prediction (NWP) model outputs; these techniques are called model output statistics (MOS) (Klein & Glahn, 1974). Generally numerical weather models, using physical and meteorological considerations, succeed in forecasting upper air patterns but are too crude to account for local variations in surface weather. Pure statistical models on the other hand, which employ only time-series data and predict future values by taking past history into account, are excellent at forecasting idiosyncrasies in local weather but are usually worthless beyond about six hours. The numerical (physical) methods have advantages in long-term prediction, while the statistical methods do well in short-term forecasts. The MOS search statistical relations between the forecast meteorological fields and the observed parameters, thus it combines two procedures: complex numerical forecasts based on the physics of the atmosphere modeling the large-scale weather patterns and regression equations in statistical post-processing to clarify surface weather details (Vannitsem, 2008). The MOS takes into account local defects and biases (systematic errors) of the deterministic model, which may arise for many reasons including the inability to account for physical processes at a scale smaller than the grid used in the numerical solution of the model equations. The objective of bias corrections is to minimize the systematic errors of the next forecast using bias from past errors, which result from many sources in NWP modeling systems, e.g. the physical parameterization of weather events. The local adaptation of weather forecasts may be done by means of “Perfect prog” method, which determines statistical relations between grid point values of meteorological field analyses and observations. Kalman filtering is used to recalculate the mean values and variances of the regression coefficients for the forecast parameters obtained by linear regression taking into account the differences between the parameter last predicted and observed values (Coiffier, 2011). Other statistical algorithms used for minimizing the bias of the next forecast may apply “Running-mean” bias corrections and the “Nearest neighborhood weighted mean” (Durai & Bhradwaj, 2014). The

* Tel.: +420 5973297970; fax: +420 597329797.
E-mail address: lzjavka@gmail.com

http://dx.doi.org/10.1016/j.eswa.2015.08.057
0957-4174/

hybrid models use additional input data, starting from a given global forecast model, which does not include local characteristics. Two different down-scaling methods using an unbiased observing system (independent from the NWP model to prevent its drift) and meso-scale model, which results from the global prediction system, may be considered. A physical down-scaling applies additional satellite atmospheric measurements (Cucurull, Anthes, & Tsao, 2014) while the statistical down-scaling uses only independent surface data observations, e.g. Bayesian hierarchical approach can combine information from short and long observational records (Nott, Dunsmuir, Kohn, & Woodcock, 2001).

Further correction methods were developed to estimate and eliminate global weather system model errors, induced due to uncertain initial conditions, data and computational limitations (Danforth, Kalnay, & Miyoshi, 2007). The "Variational" method, based on previous numerical forecasts, estimates a non-systematic component of numerical forecast error. This method assumes the component is linearly dependent on some combination of the initial fields, end time forecasts and the forecast tendency, which can identify the forecasting error (Shao, Xi, & Qiu, 2009). The NWP model error can be expressed solving the inverse problem of the Lagrange interpolation polynomial, which coefficients are determined by past model performance. The length of past multi-time data observations sufficient for the optimal error estimation together with numerical model outputs and observation errors determine the method accuracy (Xue, Shen, & Chou, 2013). The probabilities of certain weather events (precipitation), which occur in a local area is possible to estimate in a "germ–grain" model using non-negative least-squares approach to determine the local rain-fall intensities and a "semi-variogram" estimation technique to find the grain (cell) size (Kriesche, Hess, Reichert, & Schmidt, 2015). The proposed correction method (Section 5) does not search relations between observed and forecasted values or some other variables of the NWP model to describe its systematic or initial output errors but it forms a general differential equation model of the changeable up-to-date function of some relevant meteorological variables from the observations, which feature the current local weather specific condition, valid for the NWP model outcomes that enter the correction model to recalculate (revise) one target prognosis series. Neural networks, trained for local fluctuating weather data relations of some few past days, can apply the same NWP output types instead of the real values (which are not known yet) to stand for the mentioned statistical local adaptations in the majority of cases.

Artificial neural networks (ANN) are able to model the non-linear nature of dynamic processes and reproduce an empirical relationship between some inputs and one or more outputs. A common ANN operating principle is based on learned entire similarity relationships between new presented input patterns and the trained ones. It does not allow for eventual direct elementary data relations, which multi-variable low-order polynomial functions can easy describe (Nikolaev & Iba, 2006). Differential equations can model to advantage physical or natural dynamic systems, which can be hardly or only with difficulties described by means of unique explicit functions; the solutions can apply power (Balser, 2004) or wave series (Chaquet & Carmona, 2012) and ANN structure to substitute for predetermined differential equations (He, Reif, & Unbehauen, 2000; Jianyu, Siwei, Yingjian, & Yaping, 2003). Extended polynomial networks may apply some mathematical principles to define and solve general differential equations. It is possible to express a general connection between input and output variables by means of the Volterra functional series, a discrete analogue of which is the Kolmogorov–Gabor polynomial (1).

$$y = a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ijk} x_i x_j x_k + \ldots$$
$$n - number\ of\ variables X(x_1, x_2, \ldots, x_n) \quad A(a_1, a_2, \ldots, a_n), \ldots - vectors\ of\ parameters \tag{1}$$

Group Method of Data Handling (GMDH) was created by a Ukrainian scientist Aleksey Ivakhnenko in 1968, when the back-propagation technique was not known yet. It forms a multi-layer polynomial neural network (PNN) in successive steps, adding one layer a time, which decomposes the complexity of the process, expressed by the Kolmogorov–Gabor polynomial, into many simpler relationships, each described by the low order polynomial transfer function (2) for every pair of the input values x_i, x_j (Ivakhnenko, 1971).

$$y = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_j^2 + a_5 x_j^2 \tag{2}$$

Differential polynomial neural network (D-PNN) is a new neural network type, designed by the author, which extends the basic GMDH-PNN structure; however its operating and constructing principles differ from those of the GMDH, based on Taylor-series expansions. D-PNN generates convergent sum series of relative polynomial derivative terms, which together form and substitute for an unknown general partial differential equation (DE) solution of the approximation of a complex function, described by data observations. The derivative series model is in principle quite different from direct composing computational techniques, which use a collection of operators and terminals of a predefined set to form symbolic tree-like structural expressions. The genetic programming can solve explicit DE forms (Cornforth & Lipson, 2013) or a pre-defined general DE system (Iba, 2008) after this manner. The D-PNN can combine the PNN functionality with some mathematical principles of DE substitutions. Its models lie on the boundary of neural networks and exact computational techniques. The D-PNN relative data processing is based on the derivative polynomial generalization of data relations in a searched function model, which composite selective series descriptions facilitate a much larger variety of model forms than usual and that allow apply a wider range of test input or output values than defined by a training data set (Zjavka, 2013a). The general partial DE is decomposed into a multi-layer network structure to produce substitution sum series for the derivative model, the exact solution of which is problematic or impossible to get using direct composing techniques or the DE explicit form is unknown (Zjavka, 2013a).

2. General partial differential equation composition

D-PNN forms and solves the general partial DE (3), in which an exact definition is not known in advance and which can generally describe a system model using summation derivative terms. The searched function u may be expressed in the form of sum series (4), consisting of series arising from derivative sum convergent term series (5) in the case of 2 input variables. The study substitutes the general partial DE (3) with multi-variable polynomial fraction terms (8), which can describe partial relative derivative changes of some input variables combinations in a sum series solution. The simple form of an unknown searched u function is possible to calculate from the general DE (3) as the sum of the rest of the derivative terms, i.e. its partial derivatives (4).

$$a + bu + \sum_{i=1}^{n} c_i \frac{\partial u}{\partial x_i} + \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \ldots = 0$$
$$a, b, \mathbf{c}(c_1, c_2, \ldots, c_n), \mathbf{d}(d_{11}, dc_{12,}, \ldots), \ldots - parameters$$
$$x(x_1, x_2, \ldots, x_n) - vector\ of\ n - input\ variables \tag{3}$$

$$u = \sum_{k=1}^{\infty} u_k$$
$$n - number\ of\ input\ variables$$
$$u(\mathbf{x}) - unknown\ function\ of\ n - variables \tag{4}$$