ARTIFICIAL INTELLIGENCE APPROACH OF MODELING OF PM10 EMISSION CLOSE TO A STEEL PLANT

Miha Kovačič
ŠTORE STEEL d.o.o.
Štore
Laboratory for Multiphase Processes, University of Nova Gorica
Nova Gorica
Slovenia

Sandra Senčič
KOVA d.o.o.
Celje
Slovenia

Uroš Župerl
Faculty of mechanical engineering, University of Maribor
Maribor
Slovenia

ABSTRACT
To implement sound air quality policies, regulatory agencies require tools to evaluate the outcomes and costs associated with various emission reduction strategies. The applicability of such tools can also remain uncertain. It is furthermore known that source-receptor models cannot be implemented through deterministic modeling. The article presents an attempt of PM10 emission modeling carried close to a steel production area with the genetic programming method. The daily PM10 concentrations, daily rolling mill and steel plant production, meteorological data (wind speed and direction – hourly average, air temperature – hourly average and rainfall – daily average), weekday and month number were used for modeling during a monitoring campaign of almost half a year (23.6.2010 to 12.12.2010). The genetic programming modeling results show good agreement with measured daily PM10 concentrations.

Keywords: Steel plant, PM10 concentrations, modeling, genetic programming

1. INTRODUCTION
Particulate matter (PM) pollution is, especially in residential areas near industrial areas, a problem of great concern. This is not only because of the adverse health effects but also because of reduced visibility; on a global scale, effects on the radiative balance are also of great importance [1-3].
To reduce PM levels in the air a deep knowledge of the contributing sources, background emissions, the influence of the meteorological conditions, as well as of PM10 formation and transport processes is needed.

However, current state-of-the-art PM10 modeling does not allow us to quantitatively model the whole range of emissions behavior, which is why the dispersion modeling is thus increasingly connected with intelligent algorithms such as artificial neural networks [4-9] and evolutionary computation [9].
The objective of this work was to model PM10 emissions close to a steel plant area in Slovenia by means of a genetic programming and artificial network method. Genetic programming method has been proven to be an effective optimization tool for multicriterial and multiparametrical problems [10-13].

2. EXPERIMENTAL SETUP

Figure 1 shows the locations of the sampling sites, rolling mill, steel plant and residential areas. Due to rolling mill and steel plant PM10 contribution also several source categories influence the PM10 concentrations. These include combustion and non-combustion traffic sources, urban background concentrations, along with both contributions that are transported by regionally and long-range.

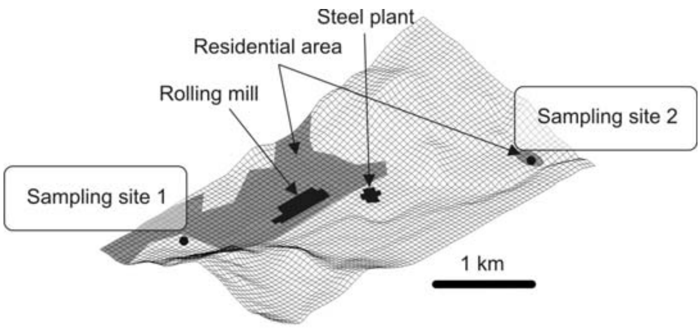

Figure 1. Topographic view of the study area

Samples for this study were collected between 23.6.2010 and 12.12.2010. Sampling was performed 1.5 m above the ground. PM10 samples were collected for 24 hours on Mondays using low-volume samplers equipped with EPA-equivalent size-selective inlets. Particles with diameter 10 μm (PM10) were collected on cellulose esters membranes with high collection efficiencies (99%). In total 172 PM10 samples for each sampling site were available.
Hourly average air temperature, wind speed and direction and daily rainfall data were made available to the authors by the Slovenian Environment Agency.
The hourly data based on electric arc and rolling mill production was collected during the study period. During the study period, the electric arc furnace was stopped for 28465 minutes and the rolling mill was stopped for 8213 minutes.

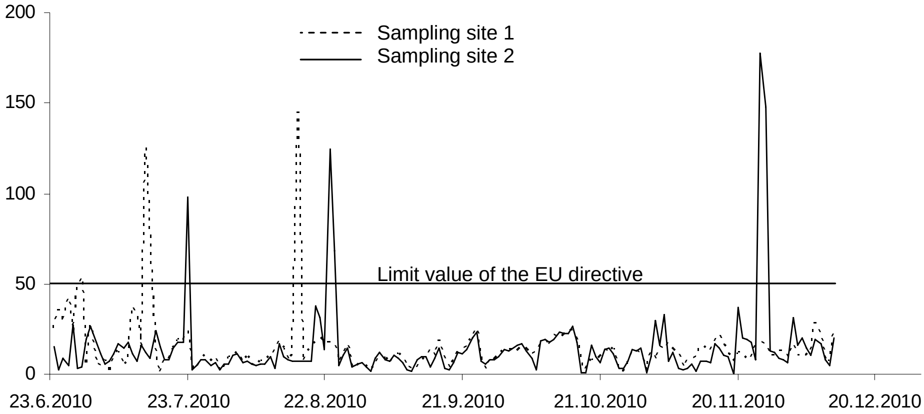

Figure 2. The measured PM10 concentrations during the study period for the sampling sites

3. GENETIC PROGRAMMING MODELING

Genetic programming is probably the most general evolutionary optimization method [13]. The organisms that undergo adaptation are in fact mathematical expressions (models) for the PM10 concentrations prediction in the present work. The concentration prediction is based on the available function genes (i.e., basic arithmetical functions) and terminal genes (i.e., independent input parameters, and random floating-point constants). In the present case the models consist of the following function genes: addition (+), subtraction (-), multiplication (*) and division (/), and the following terminal genes: weekday (*WEEKDAY*) and month number (*MONTH*), wind speed [m/s] (*SPEED*), wind direction [°] (*DIRECTION*), air temperature [°C] (*TEMP*), rainfall [ml] (*RAIN*), electro arc furnace efficiency [min/hour] (*EAF*), rolling mill efficiency [min/hour] (*ROLLING*). In order to ascertain the influence of seasons and traffic during workday hours the weekday and month number were also added as terminal genes.

Random computer programs of various forms and lengths are generated by means of the selected genes at the beginning of the simulated evolution. The varying of the computer programs is performed by means of the genetic operations during several iterations, known as generations. After the completion of the variation of the computer programs a new generation is obtained. Each generation is compared with the experimental data. The process of changing and evaluating organisms is repeated until the termination criterion of the process is fulfilled. The maximum number of generations is chosen as a termination criterion in the present algorithm.

The following evolutionary parameters were selected for the process of simulated evolutions: 500 for the size of the population of organisms, 100 for the maximum number of generations, 0.4 for the reproduction probability, 0.6 for the crossover probability, 6 for the maximum permissible depth in the creation of the population, 10 for the maximum permissible depth after the operation of crossover of two organisms, and 2 for the smallest permissible depth of organisms in generating new organisms. Genetic operations of reproduction and crossover were used. For selection of organisms the tournament method with tournament size 7 was used. 100 independent civilizations of mathematical models for prediction of the PM10 concentration were developed. The best evolution sequence of 100 generations was computed in 8 hours and 41 minutes on 2.39 GHz processor and 2 GB of RAM by an AutoLISP based in-house coded computer program.

The model fitness f has been defined as:

$$f = \sum_{i=1}^{n} \left(P_i - M_i \right) + N \cdot 10000 \quad , \tag{1}$$

where n is the size of sample data and, P_i is predicted PM10 concentration, M_i is measured PM10 concentration and N is the number of all cases when:

$$P_i < 50 \wedge M_i > 50 \vee M_i < 50 \wedge P_i > 50 \quad . \tag{2}$$

The limit value of the EU directive, i.e. a daily mean PM10 concentration, is 50 μg/m^3.

The simulated evolution in one run of the genetic programming system (out of 100) produced the following best model for prediction of PM10 concentration for sampling site 1:

$$\begin{gathered}
\Bigg(\text{DIRECTION} + \Bigg(\text{DIRECTION} + \text{MONTH} + \text{WEEKDAY} + \frac{\text{SPEED}\,(-1.99563 + 3\,\text{WEEKDAY} + 4\,\text{MONTH WEEKDAY})}{\text{MONTH} - \text{WEEKDAY}} + \\
\frac{\text{WEEKDAY}\,(-1.99563 + \text{MONTH WEEKDAY})}{\text{MONTH} - \frac{\frac{1}{2} + \text{MONTH WEEKDAY}}{\text{MONTH}}}\Bigg) \Bigg/ \Bigg(\frac{-1.99563 + \text{MONTH WEEKDAY}}{\text{MONTH} + \text{WEEKDAY} - \frac{-1.99563 + \text{MONTH WEEKDAY}}{\text{MONTH} - \frac{\text{WEEKDAY}}{\text{SPEED}}}} + \\
\Bigg((-1.99563 + \text{MONTH WEEKDAY})\Bigg(-\text{MONTH} + (\text{DIRECTION} - \text{WEEKDAY}) \Bigg/ \Bigg(\text{DIRECTION} + \\
\text{WEEKDAY} + \text{MONTH WEEKDAY} - \frac{\text{MONTH SPEED}\,(-1.99563 + \text{MONTH WEEKDAY})}{-\text{DIRECTION} + \text{MONTH} + \text{WEEKDAY}} - \\
\frac{\text{MONTH}^2\,(-1.99563 + \text{MONTH WEEKDAY})}{\text{MONTH} - \frac{\text{WEEKDAY}}{\text{DIRECTION}}}\Bigg)\Bigg)\Bigg) \Bigg/ \\
\Bigg(\text{MONTH} - \frac{2\,\text{WEEKDAY} + \text{MONTH WEEKDAY}}{\text{DIRECTION WEEKDAY}}\Bigg)\Bigg)\Bigg) \Bigg/ \Bigg(\text{DIRECTION} - \text{RAIN} + \\
\left(\text{MONTH} - \frac{\text{MONTH} - \text{WEEKDAY}}{\text{WEEKDAY}}\right) + \text{MONTH WEEKDAY} - \\
\left(-\text{MONTH} + \left(\text{MONTH} - \frac{\text{MONTH} - \text{WEEKDAY}}{\text{SPEED}}\right)\right) \\
(-1.99563 + \text{MONTH WEEKDAY})\Bigg)
\end{gathered} \quad , \tag{3}$$

with fitness of 1019.95, number $N = 0$ and average deviation of 5.96 μg/m^3.

The best evolutionary developed model (out of 100) for prediction of PM10 concentration for sampling site 2 is:

$$\frac{\left(-5.76596+\mathrm{MONTH}+\frac{\mathrm{TEMP}}{\mathrm{DIRECTION}}\right)\left(\frac{-\frac{\mathrm{SPEED}}{\mathrm{TEMP}}+\frac{1-\frac{\mathrm{SPEED}}{\mathrm{TEMP}}}{\mathrm{WEEKDAY}}}{\mathrm{MONTH}}+\frac{\left(\mathrm{EAF}+\frac{\mathrm{MONTH}+\frac{\mathrm{TEMP}}{\mathrm{DIRECTION}}-\mathrm{WEEKDAY}}{\mathrm{EAF}+\frac{\mathrm{SPEED}}{\mathrm{DIRECTION\ TEMP}}+\frac{\mathrm{TEMP}}{\mathrm{EAF}}}\right)\mathrm{WEEKDAY}}{\mathrm{TEMP}\left(\mathrm{TEMP}+\frac{-\frac{\mathrm{SPEED}}{\mathrm{TEMP}}+\frac{1-\frac{\mathrm{SPEED}}{\mathrm{TEMP}}}{\mathrm{WEEKDAY}}}{\mathrm{DIRECTION}}+\frac{\mathrm{EAF\ RAIN\ SPEED}}{\mathrm{TEMP}\left(-\mathrm{TEMP}+\frac{\mathrm{TEMP}}{\mathrm{WEEKDAY}}\right)\mathrm{WEEKDAY}}\right)}\right)}{\mathrm{DIRECTION\ WEEKDAY}}+$$

$$+\frac{\mathrm{MONTH}-\frac{\mathrm{WEEKDAY}^2}{\mathrm{TEMP}}+\frac{\mathrm{SPEED\ TEMP}^2}{\mathrm{WEEKDAY}\left(\mathrm{DIRECTION}+\frac{\mathrm{TEMP}\,(-\mathrm{TEMP}+\mathrm{WEEKDAY})}{\mathrm{DIRECTION}}\right)}}{\mathrm{DIRECTION}+\frac{\mathrm{MONTH}-\mathrm{WEEKDAY}+\frac{\mathrm{TEMP}+\frac{\mathrm{EAF\ WEEKDAY}}{\mathrm{SPEED\ TEMP}}+\frac{-\mathrm{TEMP}+\mathrm{WEEKDAY}}{\mathrm{DIRECTION}}}{\mathrm{DIRECTION}}}{\mathrm{EAF}+\frac{\mathrm{EAF\ RAIN}}{\mathrm{TEMP\ WEEKDAY}}+\frac{-\mathrm{TEMP}+\mathrm{WEEKDAY}+\frac{-\mathrm{TEMP}+\mathrm{WEEKDAY}}{\mathrm{DIRECTION}}}{\mathrm{DIRECTION}}}}+0.30872\quad,\qquad(4)$$

with fitness of 11124.67, number $N = 1$ (on the 30.6.2010 the measured PM10 concentrations were 53.6 $\mu g/m^3$ and predicted 21.41 $\mu g/m^3$), and average deviation of 6.54 $\mu g/m^3$.

4. CONCLUSIONS

This paper presented the possibility of the PM10 concentration prediction close to a steel plant area with genetic programming. The daily PM10 concentrations, daily rolling mill and steel plant production, meteorological data (wind speed and direction – hourly average, air temperature – hourly average and rainfall – daily average), weekday and month number were used for modeling during a monitoring campaign of almost half a year (23.6.2010 to 12.12.2010). The special fitness function for genetic programming system was designed in order to assure also PM10 limit value exceedance prediction. For each sampling site the best models for PM10 prediction were obtained from 100 runs of the genetic programming system. The model for sampling sites 1 and 2 predicts concentrations within an average error range of 5.96 $\mu g/m^3$ and 6.54 $\mu g/m^3$, respectively. All exceedances of the EU directive limit value (50 $\mu g/m^3$) were administered at sampling site 1, but only 4 out of 5 of these occurred at sampling site 2. The number, when the prediction is above that limit value, when in order to assure PM10 concentration exceedance prediction by developed predictive model it should in fact be below the limit, and also when prediction by developed predictive model should be above the limit, was 1 at sampling site 1 and 0 at sampling site 2. In the future we will carry out genetic programming based dispersion modeling according to the calculated wind field, air temperature, humidity and rainfall in a 3D Cartesian coordinate system. The prospects for arriving at a robust and faster alternative to the well-known Lagrangian and Gaussian dispersion models are optimistic.

5. REFERENCES

[1] Marcazzan, G. M., Ceriani, M., Valli, G., Vecchi, R.: Source apportionment of PM10 and PM2.5 in Milan (Italy) using receptor modeling, The Science of the Total Environment, 2003, 317, 137-147.

[2] Watson, J. G.: Visibility: science and regulation, Journal of the Air and Waste Management Association, 2002, 52, 628-713.

[3] Vrins, E., Schofield, N.: Fugitive dust emission by an ironmaking site, Journal of Aerosol Science, 2000, 31, 524-525.
[4] Kukkonena, J., Partanena, L., Karppinena, A., Ruuskanenb, J., Junninenb, H., Kolehmainenb, M., Niskab, H., Dorlingc, S., Chattertonc, T., Foxalld, R., Cawleyd, G.: Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki, Atmospheric Environment, 2003, 37, 4539-4550.
[5] Zhou, H., Cen, K., Fan, J.: Modeling and optimization of the NOx emission, characteristics of a tangentially fired boiler with artificial, neural networks, Energy, 2004, 29, 167-183.
[6] Hooyberghsa, J., Mensinka, C., Dumontb, G., Fierensb, F., Brasseurc, O.: A neural network forecast for daily average PM10 concentrations in Belgium, Atmospheric Environment, 2005, 39, 3279-3289.
[7] Perez, P., Reyes, J.: An integrated neural network model for PM10 forecasting, Atmospheric Environment, 2006, 40, 2845-2851.
[8] Grivas, G., Chaloulakou, A.: Artificial neural network models for prediction of PM10 hourly concentrations, in the Greater Area of Athens, Greece, Atmospheric Environment, 2006, 40, 1216–1229.
[9] Kovačič, M., Uratnik, P., Brezočnik, M., Turk, R.: Prediction of the bending capability of rolled metal sheet by genetic programming, Materials and Manufacturing Processes, 2007, 22, 634-640.
[10] Kovačič, M., Šarler, B.: Application of the genetic programming for increasing the soft annealing productivity in steel industry, Materials and Manufacturing Processes, 2009, 24(3), 369-374.
[11] Kovačič, M.: Genetic programming and Jominy test modeling, Materials and Manufacturing Processes 2009, 24(7), 806-808.
[12] Kovačič, M., Senčič, S.: Critical inclusion size in spring steel and genetic programming, RMZ - Materials and Geoenvironment, 2010, 57(1), 17-23.
[13] Koza, JR.: Genetic programming III, Morgan Kaufmann, San Francisco, 1999.