Turkish Journal of Physics

http://journals.tubitak.gov.tr/physics/

Research Article

Turk J Phys
(2018) 42: 273 – 282

doi:10.3906/fiz-1712-14

Function mining for the $<P_T>$-N_{ch} correlations in pp and pp(bar) collisions based on symbolic regression

EL-Sayed El-DAHSHAN[1,2,*]
[1]Department of Physics, Faculty of Sciences, Ain Shams University, Cairo, Egypt
[2]Egyptian E-Learning University (EELU), El-Geiza, Egypt

Received: 24.12.2017 • **Accepted/Published Online:** 11.05.2018 • **Final Version:** 01.06.2018

Abstract: The investigation and analysis of the correlation between the mean transverse momentum ($< P_T >$) of charged particles and the charged particle multiplicity (N_{ch}) allow physicists to understand the contribution of multiple-parton interactions to the particle production mechanism. A symbolic regression (SR) method, based on gene expression programming (GEP), is proposed for mining a function that describes the $< P_T >$-N_{ch} correlation in proton–proton and proton–antiproton (pp and pp(bar)) collisions at collision energies from IRS to LHC. The discovered function simulates and models the correlation between $< P_T >$ and N_{ch} in wide energy range $s^{1/2}$. In the framework of the proposed GEP model for $< P_T > -N_{ch}$ correlation, the equation obtained describes the main features of the experimental data. Predictions for $< P_T > -N_{ch}$ correlations at the future LHC collision energy of 14 TeV are obtained. The accuracy of the calculated and predicted results is assessed by comparing them with the available experimental data and the theoretical ones.

Key words: Modeling and simulation, pp(bar) and pp collisions, charged multiplicity, transverse momentum, correlations, symbolic regression, function mining

1. Introduction

The role of scientific prediction and modeling of the fundamental nature of matter and atom began about 2500 years ago with ancient Greek philosophers with the Democritus (460–370 BC) atom model “All matter is made of indivisible particles called atoms”. In the early years of the 20th century, Rutherford [1] and Bohr [2] predicted and modeled the subatomic structure. The fundamental particles and the forces that govern their interactions were predicted by the Standard Model (SM) throughout the latter half of the 20th century. Unfortunately, the SM is not a complete theory. Furthermore, string theory and super-symmetry (if you combine super-symmetry with string theory, you have superstring theory) have their limitations [3].

The power of a computational model is that it allows scientists to simulate and model the mechanism and behavior of complex and nonlinear systems by the use of mathematics, as well as statistical and computer science. The application of computers in simulation and modeling physics began with the work by Feynman with the first use of computer science in physics modeling [4].

Today, the applications of computer simulation, based on artificial intelligence, have drawn a momentous amount of attention and interest for researchers as well as scientists. This is due to their potential for

*Correspondence: e_eldahshan@yahoo.com

various applications in classification, regression, and function discovering and mining [5,6], such as neural networks (NNs), evolutionary computing (EC) algorithms (e.g., genetic programming (GP) and gene expression programming (GEP)) and support vector machines (SVMs) [7–9]. They have been widely utilized by many researchers in various areas of physics and other sciences (e.g., Teodorescu (2009), El-Dahshan et al. (2008, 2009), Link et al. (2005)) [10–13].

The development of computational paradigms (based on symbolic regression (SR)) [14,15] in physics, describing and predicting the experimental data of the high energy particles interactions, has long been regarded as an important and widely studied issue in the academic and research community. This kind of computational paradigm (SR) has led to the discovery of an empirical equation that describes the current experiments, serves as a guideline in designing new experiments, and allows physicists to carry out stronger tests of many theoretical conjectures (estimations) [14,15].

Recently, the application of symbolic regression via EC such as GP [9] and GEP [16] has been widely used by many researchers in different areas of physics. The goal of the SR is to discover (mine) symbolic functions (the symbolic function consists of a function set, e.g., sin, cos, and log, and a terminal set, e.g., independent variables and random constants), which can be complicated in terms of its structure, in the correct simple form, and find appreciate numeric coefficients that approximate the given experimental data set [16]. SR paradigms have become popular in physics data analysis and predict results of future experiments. They are used to perform complex nonlinear regression by intelligent pattern recognition. This kind of artificial intelligence is achieved by the principle "learning from examples" [14,15,17].

The problem of finding a predictive function (function discovering and mining), based on experimental data, has become an interesting issue in physics data analysis [14,17]. In order to improve the intelligibility of function approximation and discovery in the form of a symbolic function, the symbolic regression method is used [14–16].

The application of SR paradigms for function mining is motivated by the lack of knowledge about the underlying physical phenomena. Learning from examples is the only possibility, since no theory is available that would allow building an algorithm in the classical way. However, it has the disadvantage of not giving an explanation of the subjacent mechanisms.

Many new physics results will be presented as a direct product from the successful application of different SR methods in different physics and science experiments to model and predict the experimental data; for example, Derouich et al. [18] applied it in astronomy and astrophysics, Schmidt and Lipson [17] used it to rediscover the fundamental laws of nature, Cai [19] used it to model heat transfer correlations, Quade [20] investigated the prediction of dynamic systems using symbolic regression, Hills et al. [21] discovered Lagrangians automatically from data based on SR, Kurse et al. [22] studied the extrapolatable analytical functions using this technique, Yang [23] modeled oil production based on SR, Murari [24] investigated the derivation of confinement scaling laws, Golafshania [25] predicted the self-compacting concrete elastic modulus using two SR techniques, Guven and Aytek [26] built a new approach for modeling the stage-discharge relationship in hydrology, and Guven et al. [27] empirically modeled evapotranspiration. The success of the SR methods in the above-mentioned fields encourages us to use it in modeling and investigating the behavior of pp scattering.

The study of the correlation between the mean transverse momentum "$\langle P_T \rangle$" of charged particles and the charged particle multiplicity "N_{ch}" (i.e. $\langle P_T \rangle - N_{ch}$ correlation) provides information about both soft and hard scattering and it is sensitive to the modeling of multiple-parton interactions [28–30].

The $\langle P_T \rangle - N_{ch}$ correlation, which is first observed at $Sp\bar{p}S$ and ISR colliders, has been studied by many experiments at hadron colliders in pp and $p\bar{p}$, covering collision energies from (ISR) $\sqrt{s} = 19$ GeV up to 13 TeV (the recent LHC experiment). The increase in P_T with N_{ch} in the central rapidity region observed in all experiments can be modeled and reproduced in the PYTHIA event generator [29–36].

In the present study, the EC symbolic regression technique employed is the GEP [16]. The choice of the GEP approach is due to its ability to form a mathematical model (symbolic function) directly from the available experimental data to analyze the given phenomena that describes the current experiments and serves as a guideline in designing new experiments.

In the present work, the GEP method is used to build a mathematical model to calculate and predict the $\langle P_T \rangle$ as a function of N_{ch} and the center of mass energy ($\sqrt{s}$). The GEP model will directly evolve from a set of available experimental data for $\langle P_T \rangle - N_{ch}$ correlation [37–43].

The rest of the article is structured as follows. In Section 2, details of the application of SR via GEP to model the $\langle P_T \rangle - N_{ch}$ correlation and their energy dependence are given. The results obtained are presented in Section 3. Finally, Section 4 provides the findings and conclusions.

2. Symbolic regression modeling for $\langle P_T \rangle - N_{ch}$ obtained by GEP

The aim of this study is to mine and develop a function to simulate and model the $\langle P_T \rangle - N_{ch}$ correlation and their energy dependence in pp and $p\bar{p}$ collision. To achieve this aim, gene expression programming (GEP), which promises an evolutionary algorithm of symbolic regression techniques, is used. The objective of the SR modeling problem is to discover a suitable mathematical model that can approximately express the behavior of a nonlinear complex system directly from experimental data [15–17].

The SR is contrary to the traditional regression technique, which searches for the best parameters of a presumed regression function that minimize the difference between the observed and the calculated values, while SR infers the model (symbolic function) from a data search for finding both the best structure and parameters of a model for which no explicit equation exists. The SR can be used to find the approximation solution to a problem via some evolutionary algorithm, such as GP and GEP. The GEP algorithm can discover a simpler and more explicit mathematical expression (develop both the structure and parameters) for describing the given experimental data [37–45].

In this paper, the developed GEP mathematical model for the $\langle P_T \rangle = \bar{P}(\sqrt{s}, N_{ch})$ is given. The GEP model developed herein is mainly aimed at generating mathematical functions for the calculation and prediction of $pp(\bar{p})$ interactions $\bar{P}(\sqrt{s}, N_{ch})$ for different high and ultrahigh energies (up to 14 TeV).

DTREG software [46] is used for the GEP model, using a k-fold cross validation [7] technique for producing highly accurate results and assessing the accuracy of the model without requiring an independent test dataset.

In order to find out how accurate the results and the uncertainty of predictions of the developed GEP model are, some statistical verification criteria are used, such as squared Pearson correlation coefficient (CC), mean absolute error (MAE), and normalized-mean-square error (NMSE) [47]. The NMSE and MAE are used to test the deviation of the model calculations and prediction from the actual data (experimental data). The CC is used to measure to what extent the model predicts data obeying the trend of the experimental one.

2.1. Overview of gene expression programming (GEP)

GEP is introduced by Ferreira [16]. It is an evolutionary computational technique similar to GA and GP approaches, which are very successful and powerful in solving SR problems. These approaches (GA, GP, and GEP) are inspired by Darwin's theory of natural selection [48] and the Mendelian genetic operators (such as natural operators crossover, mutation, inversion, selection) [49]. The GEP algorithm can discover a simpler and more explicit mathematical expression (develop both the structure and parameters) for describing the given data in the form of a symbolic function. GEP evolves a computer program to solve the given problem. The programs are represented in chromosomes. GEP works by creating a population, consisting of individuals (chromosomes). Each individual represents a possible solution for the given problem. The chromosome "genotype" is then translated into an expression tree (ET) "phenotype" (through genotype–phenotype mapping). A chromosome is composed of one or more genes, and each gene consists of two parts: a head (contains symbols from functions or terminals) and a tail (contains symbols from terminals set). Genetic operations (crossover, mutation, inversion, selection) are applied to a chromosome (individual) to create a new population that keeps the fittest member of the previous generation based on the fitness function [8,9]. During creation of a new individual program, other operators like mutation and selection are applied. The individual is then evaluated and compared to the stopping criteria (mean square error (MSE)). This process is repeated for a prespecified number of generations or until a solution is found. The methodology to evolve the population that forms the basis for this work is depicted in Figure 1. In Figure 1 an example for the encoding of a chromosome as a linear string with two genes and corresponding ET is indicated. The mathematical equation represented by this chromosome is also shown in Figure 1.

The GEP algorithm is used to discover (mining) a symbolic function that accurately describes the correlation $\langle P_T \rangle - N_{ch}$. The GEP model developed herein is mainly aimed at generating the mathematical functions for the simulation and modeling of $\langle P_T \rangle - N_{ch}$ correlation at different $\sqrt{s}$. In this study the GEP developed model has 2 inputs (N_{ch} and $\sqrt{s}$) and one output $\bar{P}(\sqrt{s}, N_{ch})$.

The discovered function enables the prediction of $\langle P_T \rangle - N_{ch}$ correlation of new roots that are outside the available experimental data [37–45], as well as facilitation of human insight and understatement of the dependence of $\langle P_T \rangle$ on the N_{ch} and $\sqrt{s}$.

After several trials, the optimum GEP parameters giving the best function (Eq. (1)) that describes the $\langle P_T \rangle - N_{ch}$ as a function of N_{ch} and $\sqrt{s}$ are obtained and summarized as follows:

Parameters of GEP model: Population size "60"; Generation "10,000"; Number of genes " 3"; Length of the gene head "13"; Max. generation "20,000"; Linking function "+"; Function set "+, –, /, *, cos, $\sqrt{}$, log, exp"; Mutation rate "0.044"; Recombination rate "0.3"; Inversion rate 0.4; Transposition "0.1". The uncertainty of calculations and predictions of the developed GEP model have been measured using the validation criteria: NMSE, MAE, and CC. The optimal simplified GEP model was obtained according to the validation criteria. The obtained simplified GEP model was used for calculating and predicting $<\mathbf{P_T}> = \bar{P}(\sqrt{s}, N_{ch})$ as a function of N_{ch} and $\sqrt{s}$ as given in Eq. (1).

$$\langle P_T \rangle = \bar{P}(\sqrt{s}, N_{ch}) = \left(1 + e^{-g}\right)^{-1} * \left(1 + e^{-h}\right)^{-1} * \sinh k, \tag{1}$$

where

$$g = e^{-\left[\frac{N_{ch}}{\sqrt{s}\cos(a\sqrt{s})}\right]}$$

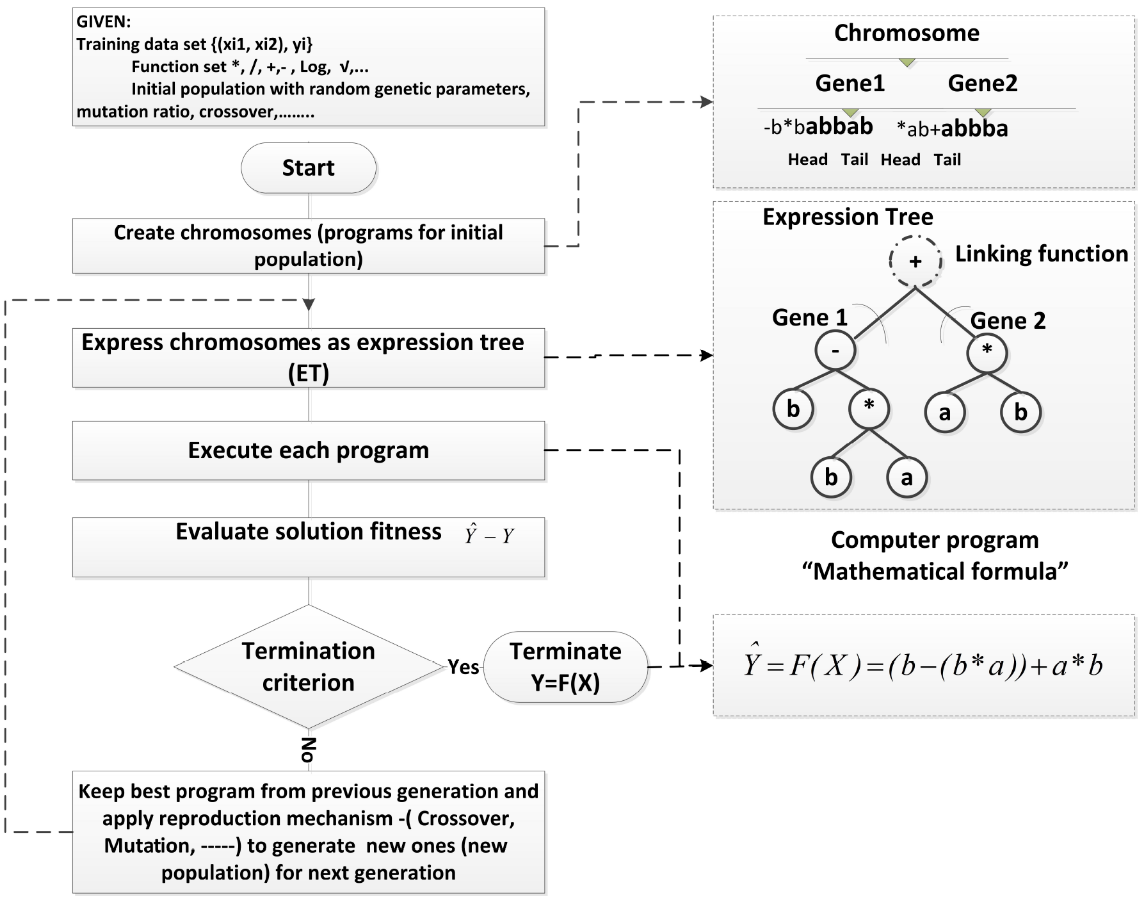

Figure 1. Flowchart of a typical GEP algorithm.

$$h = \frac{N_{ch}}{b + N_{ch} * e^{-\left(\frac{(c-\sqrt{s})^2}{d}\right)}}$$

$$k = e^{-\left[\frac{1}{1-e^{-\frac{N_{ch}}{t}}}\right]}, \quad t = m + N_{ch} * e^{-\left(\frac{(\sqrt{s}-l)^2}{q}\right)}$$

$a = 11.704, b = 11.62, c = -2593.8797, d = 203,553,969.64, m = 43.9, l = 254.89, q = 164,090.59.$

It should be noted that the proposed GEP formulations in Eq. (1) demonstrate the acceptable performance of the GEP model for estimating $\bar{P}(\sqrt{s},\, N_{ch})$ in both the training and testing stages.

3. Results and discussion

In the present work, we discovered a symbolic function based on the SR via GEP to investigate and study the correlation between $\langle P_T \rangle$ and N_{ch}. The GEP discovered function (Eq. (1)) was used to simulate the $\langle P_T \rangle - N_{Nc}$ correlation of charged particles for pp and $p\bar{p}$ interaction at $\sqrt{s} = 19$ GeV, 63 GeV, 200 GeV, 546 GeV, 900 GeV, 7 TeV, 13 TeV. The developed GEP model was tested by data sets that had not been employed in the training stage based on 10-fold cross-validation without the need for extra data.

The GEP results are compared with the experimental [37–45] data for training and testing sets. As seen from Figures 2a and 2b, the GEP model calculations agree well with the experimental data.

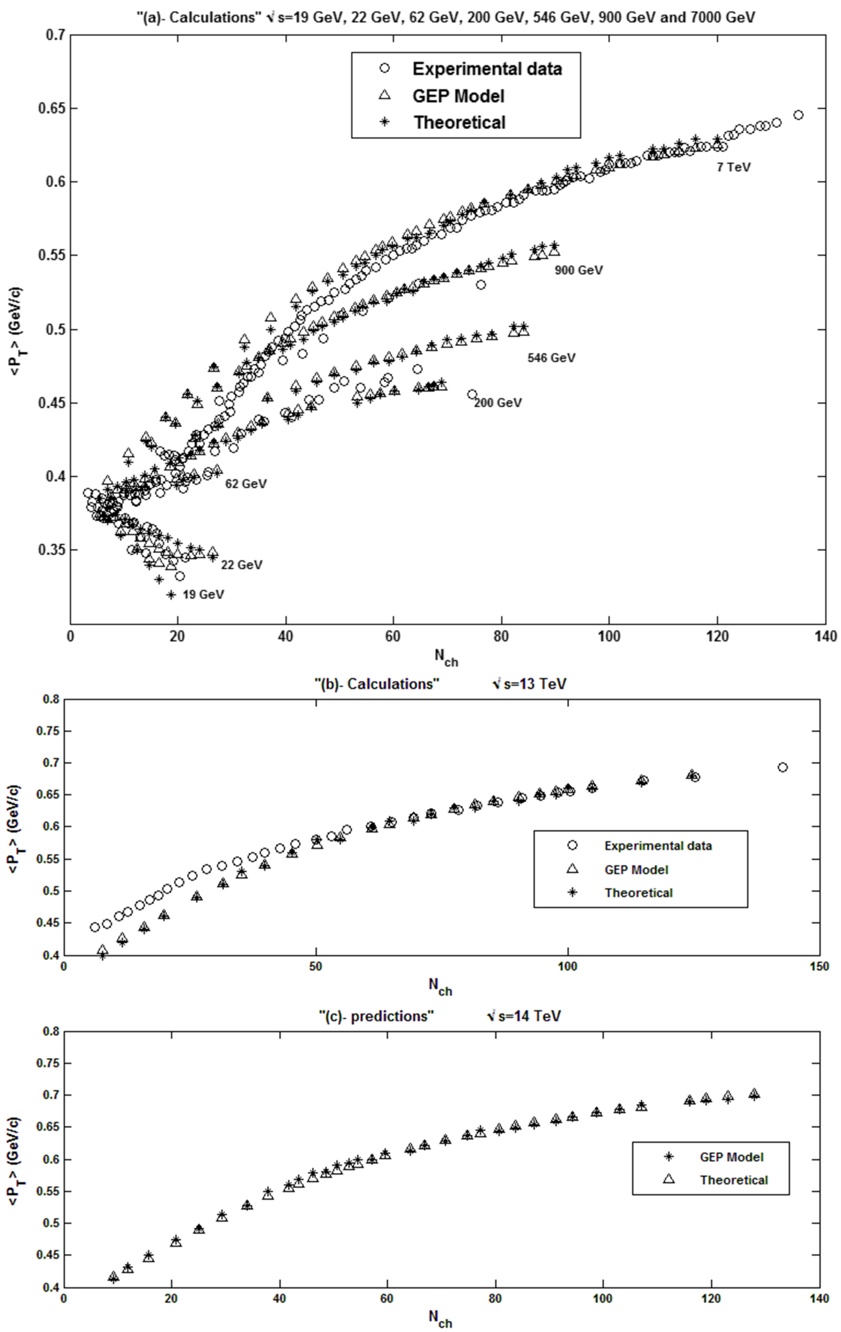

Figure 2. Comparison between our GEP approach calculated (a and b) and predicted (c) values and the corresponding experimental and theoretical ones for $\langle P_T \rangle - N_{ch}$ correlations of pp and $p\bar{p}$ interactions. (*) GEP model, Δ theoretical, and (O) experimental data.

This figure also shows that the average transverse momentum of the charged particles, $\langle P_T \rangle$, exhibits a nonzero correlation between transverse momentum and multiplicity with allowance to both positive (above 40 GeV) and negative (at low collision energies of $\sqrt{s} = 19$ GeV) $\langle P_T \rangle - N_{Nc}$ correlations. Moreover, this tendency of flattening the mean transverse momentum of charged particles with the multiplicity has been established with the growth of collision energy.

The transition from negative to positive $\langle P_T \rangle - N_{ch}$ correlation, as well as the tendency for flattening with increase of energy $\sqrt{s}$ from 19 GeV to 13,000 GeV in pp and $p\bar{p}$ collisions, is quantitatively described within our approach.

In order to evaluate the uncertainty of calculations and the performance of the GEP model, the statistical MAE, NMSE, and CC are calculated. Satisfactory agreement between the model calculated and experimental data is observed. The Table demonstrates the uncertainty of calculations and predictions (the values of CC and low MAE and NMSE) for the developed GEP model.

Table. Uncertainty of the calculations and predictions.

	MAE	NMSE	CC
Training	0.0035	0.0032	0.998
Validation	0.017	0.036	0.966

Comparing the GEP predictions with the experimental data demonstrates a high generalization capability of the proposed model. The results prove that the proposed GEP model has impressively learned well the complex relation between $\langle P_T \rangle$ and N_{ch} with high CC and low MAE and NMSE. Both our results and the other theoretical ones are within the range of experimental error. Uncertainty in the $\langle P_T \rangle - N_{ch}$ correlation measurement varies from $\pm 3\%$ at low N_{ch} to $\pm 0.5\%$ at high N_{ch} [45].

Using the discovered function (Eq. (1)), we can predict the $\bar{P}(\sqrt{s},\, N_{ch})$ values for charged particles at $\sqrt{s}$=14 TeV. The predicted values are compared with theoretical calculations [37,38] as shown in Figures 2b and 2c. From Figures 2b and 2c, we can conclude that the calculation and prediction of $\langle P_T \rangle$ versus N_{ch} correlation at 13 TeV and 14 TeV have the same trend as the theoretical one (Multi-Pomeron Exchange Model) [37,38], which supports the ability of wide usage of our models in modeling of high energy physics. The prediction of the GEP model at high energy (13 TeV, 14 TeV) is more efficient than the prediction at low energy.

The Pearson's correlation coefficient (CC) and residuals of the calculated and predicted values obtained by GEP are plotted against the experimental values of $\bar{P}(\sqrt{s},\, N_{ch})$ at different $\sqrt{s}$ as shown in Figures 3a and 3b. It can be observed from the CC and residual plot that the identified correlation by the SR methodology is nonlinear in nature for $\bar{P}(\sqrt{s},\, N_{ch})$.

The main contribution of the present work is to mine the available experimental data to discover a function that correlates the $<\mathrm{P}_T>$ and N_{ch} at different energies $\sqrt{s}$ with the use of the computational search (GEP model) without prior knowledge about the physical phenomena. The discovered GEP function that we found can help to reveal the physics underlying the observed phenomena.

4. Conclusion

The present study reports a new efficient approach for $\langle P_T \rangle - N_{ch}$ correlation modeling in pp and $p\bar{p}$ scattering. We have discovered a function that describes the correlation between $\langle P_T \rangle$ and N_{ch}. We have used a symbolic regression model to simulate and model the correlation between $\langle P_T \rangle$ and N_{ch} at various centers of mass energies $\sqrt{s}$ = 19 GeV to 13 TeV, and the prediction is performed at $\sqrt{s}$ = 14 TeV. Good agreement between our model calculations, experimental results, and other theoretical ones has been achieved. Finally, it is important to stress that the present model may serve as a robust approach and it may open a new area for the development of accurate and effective explicit formulation of $\langle P_T \rangle - N_{ch}$ correlation for future LCH energies.

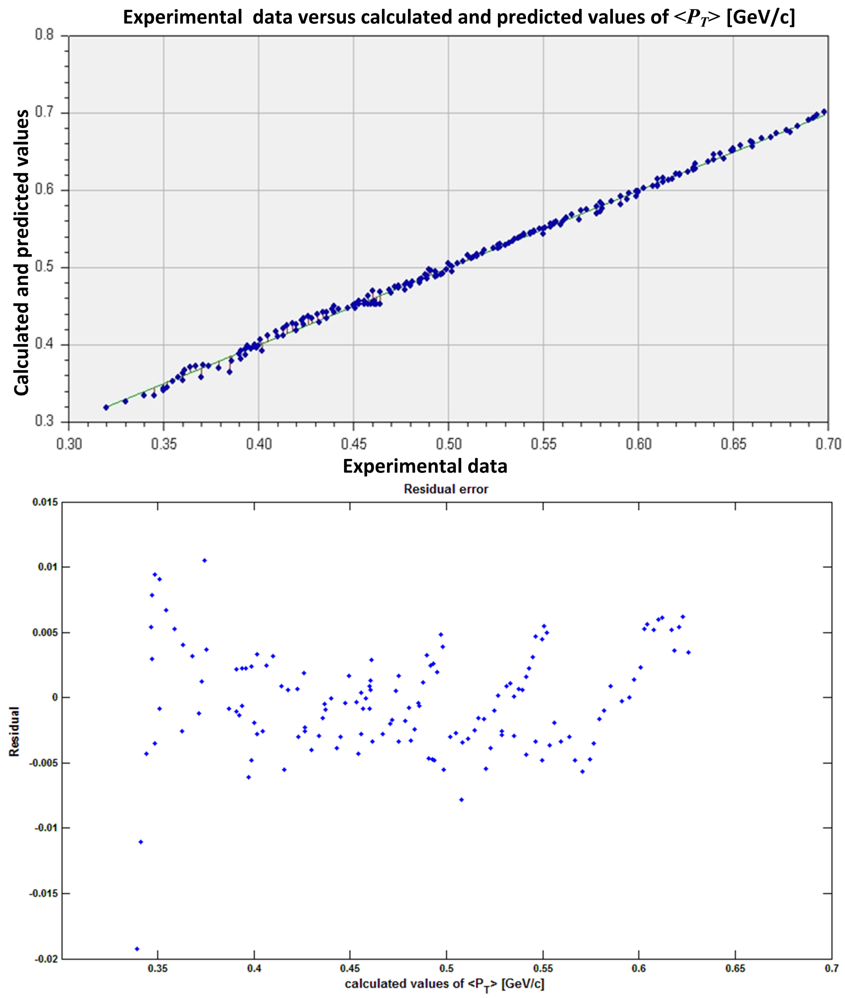

Figure 3. a. The Pearson's correlation coefficient of the experimental and predicted $<P_T>$ for the developed GEP model. b. Residual between experimental and predicted values for $\langle P_T \rangle$ versus N_{ch} for the GEP model.

References

[1] Rutherford, E. *Philos. Mag.* **1911**, *21*, 669-688.

[2] Bohr, N. *Philos. Mag.* **1913**, *25*, 10-31.

[3] Bustamante, M.; Cieri, L.; Ellis, J. *arXiv: hep-ph/0911.4409* **(2012)**.

[4] Feynman, R. P. *Int. J. Theor. Phys.* **1982**, *21*, 467-488.

[5] Džeroski, S.; Langley, P.; Todorovski, L. In *Computational Discovery of Scientific Knowledge*; Džeroski, S.; Todorovski, L., Eds. Springer: Berlin, Germany, 2007, pp. 69-97.

[6] Todorovski, L.; Dzeroski, S. In *Proceedings of the Fourteenth International Conference on Machine Learning, 1997.* San Mateo: Morgan Kaufmann Publishers Inc., 1997, p. 376.

[7] Haykin, S. *Neural Networks and Learning Machines*; Prentice Hall: Upper Saddle River, NJ, USA, 2008.

[8] Holland, J. H. *Adaptation in Natural and Artificial Systems; an Introductory Analysis with Application to Biology, Control, and Artificial Intelligence*; University of Michigan Press: Ann Arbor, MI, USA, 1975.

[9] Koza, J. R.; Keane, M. A.; Streeter, M. J.; Mydlowec, W.; Yu, J.; Lanza, G. *Genetic Programming IV: Routine Human-Competitive Machine Intelligence*, Vol. 5. Springer Science & Business Media: Berlin, Germany, 2006.

[10] Link, J. M.; Yager, P. M.; Anjos, J. C.; Bediaga, I.; Castromonte, C.; Göbel, C.; Pepe, I. M. *Nucl. Instr. Meth. Phys. Res. A* **2005**, *551*, 504-527.

[11] Teodorescu, L.; Sherwood, D. *Comp. Phys. Comm.* **2008,** *178*, 409-419.

[12] El-dahshan, E.; Radi, A.; El-Bakry, M. Y. *Int. J. Mod. Phys. C* **2009**, *20*, 1817-1825.

[13] El-dahshan, E.; Radi, A.; El-Bakry, M. Y. *Int. J. Mod. Phys. C* **2008**, *19*, 1787-1795.

[14] Schmidt, M; Lipson, H. In *Genetic Programming Theory and Practice VII, Genetic and Evolutionary Computation*; Riolo, R.; O'Reilly, U. M.; McConaghy, T. Eds. Springer: Ann Arbor, MI, USA, 2010, pp. 73-85.

[15] Kiesling, C.; Zimmermann, J. In *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004*. Proceedings of the 2004 IEEE, pp. 325.

[16] Ferreira, C. *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence (Vol. 21)*, 2nd ed.; Springer-Verlag: Berlin, Germany, 2006.

[17] Schmidt, M.; Lipson, H. *Science* **2009**, *324*, 81-85.

[18] Derouich, M.; Radi, A.; Barklem, P. S. *Astron. Astrophys.* **2015**, *584*, A64.

[19] Cai, W.; Pacheco-Vega, A.; Sen, M.; Yang, K. T. *Int. J. Heat Mass. Tran.* **2006**, *49*, 4352-4359.

[20] Quade, M.; Abel, M.; Shafi, K.; Niven, R. K.; Noack, B. R. *Phys. Rev. E*, **2016**, *94*, 012214-1-012214-15.

[21] Hills, D. J.; Grütter, A. M.; Hudson, J. J. *Peer J. Comput. Sci.* **2015**, *1*, e31.

[22] Kurse, M. U.; Lipson, H.; Valero-Cuevas, F. J. *IEEE T. Bio. Eng.* **2012**, *59*, 1572-1582.

[23] Yang, G.; Li, X.; Wang, J.; Lian, L.; Ma, T. *Energy Policy* **2015**, *82*, 48-61.

[24] Murari, A.; Peluso, E.; Gelfusa, M; Lupelli, I.; Lungaroni, M.; Gaudio, P. *Plasma Phys. Control. Fusion* **2015**, *57*, 014008.

[25] Golafshani, E. M.; Ashour, A. *Automat. Constr.* **2016**, *64*, 7-19.

[26] Guven, A.; Aytek, A. *Eng. J. Hydraul Eng.* **2009**, *14*, 812-820.

[27] Guven, A.; Aytek, A.; Yuce, M. I.; Aksoy, H. *Clean-Soil Air Water* **2008**, *36*, 905-912.

[28] Arnison, G.; Astbury, A.; Aubert, B.; Bacci, C.; Bernabei, R.; Bezaguet, A.; Böck, R.; Bowcock, T. J. V.; Calvetti, M.; Carroll, T.; et al. *Phys. Lett. B* **1982**, *118*, 167-172.

[29] Alexopoulos, T.; Allen, C.; Anderson, E. W.; Areti, H.; Banerjee, S.; Beery, P. D.; Cole, P.; Choi, Y.; De Bonte, R.; Erwin, A.; et al. *Phys. Rev. Lett.* **1988**, *60*, 1622.

[30] Breakstone, A.; Campanini, R.; Crawley, H. B.; Dallavalle, G. M.; Deninno, M. M.; Doroba, K.; Drijard, D.; Fabbri, F.; Firestone, A.; Fischer, H. G.; et al. *Phys. Lett. B* **1983**, *132*, 458-462.

[31] D'Enterria, D.; Engel, R.; Pierog, T.; Ostapchenko, S.; Werner, K. *Astropart Phys.* **2011**, *35*, 98-113.

[32] Monteno, Marco, and ALICE Collaboration, Journal of Physics: Conference Series. Vol. 5. No. 1. IOP Publishing: Bristol, UK, 2005.

[33] Bodnya, E. O.; Kovalenko, V. N.; Puchkov, A. M.; Feofilov, G. A. In *AIP Conference Proceedings*, vol. 1606, no. 1, 2014, pp. 273-282.

[34] Luettig, P.; Appelshaeuser, H.; Buesching, H.; Schuchmann, S.; Otwinowski, J. Correlation between mean transverse momentum and multiplicity in pp collisions at $\sqrt{s} = 900$ GeV in ALICE*. GSI Scientific Report 2009. 2010, p. 247.

[35] McLerran, L.; Praszalowicz, M. *arXiv: hep-ph/1006.4293* **(2010).**

[36] Bohm, G.; Zech, G. *Introduction to Statistics and Measurement Analysis for Physicists.* DESY: Hamburg Germany, 2010.

[37] Bodnia, E.; Derkach, D.; Feofilov, G.; Kovalenko, V.; Puchkov, A. *arXiv: hep-ph /1310.1627* **(2013).**

[38] Armesto, N.; Derkach, D. A.; Feofilov, G. A. *Phys. Atom. Nuclei* **2008**, *71*, 2087-2095.

[39] Aad, G.; Abbott, B.; Abdallah, J.; Abdinov, O.; Abeloos, B.; Aben, R.; Abolins, M.; Abou Zeid, O. S.; Abraham, N. L.; Abramowicz, H.; et al. ATLAS Collaboration: *Phys. Lett. B* **2016**, *758*, 67-88.

[40] Piparo, D. *Measurements of Hadron Production and Underlying Event Studies at CMS. (No. CMS-CR-2010-244).* 2010.

[41] Chatrchyan, S.; Khachatryan, V.; Sirunyan, A. M.; Tumasyan, A.; Adam, W.; Aguilo, E.; Bergauer, T.; Dragicevic, M.; Ero, J.; Fabjan, C.; et al. The CMS Collaboration, *Eur. Phys. J. C.* **2012**, *72*.10, 1-37.

[42] Khachatryan, V.; Sirunyan, A. M.; Tumasyan, A.; Adam, W.; Bergauer, T.; Dragicevic, M., Ero, J.; Fabjan, C.; Friedl, M.; Fruhwirth, R.; et al. CMS Collaboration, *Phys. Rev. Lett.* **2010**, *105*, 022002-1-022002-15.

[43] Abelev, B.; Adam, J.; Adamová, D.; Adare, A. M.; Aggarwal, M. M.; Aglieri Rinella, G.; Agnello, M.; Agocs, A. G.; Agostinelli, A.; Ahammed, Z.; et al. ALICE collaboration, *Phys. Lett. B* **2013**, *727*, 371-380.

[44] Tasevsky, M. arXiv: *hep-ex/1802.02818* (**2018**).

[45] Aad, G.; Abbott, B.; Abdallah, J.; Abdinov, O.; Abeloos, B.; Aben, R.;. Abolin, M.; AbouZeid, O. S.; Abraham, N. L.; Abramowicz, H.; Abreu, H.; et al. ATLAS collaboration, *Phys. Lett. B* **2016**, *758*, 67-88.

[46] Sherrod, P. DTREG Predictive Modeling Software (2008), Manual for software available online: www.dtreg.com.

[47] Friedrich, R.; Siegert, S.; Peinke, J.; Luck, S.; Siefert, M.; Lindemann, M. M; Raethjen, J.; Deuschl, G.; Pfister G. *Phys. Lett. A* **2000**, *271*, 217-222.

[48] Darwin, C. *On the Origin of Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life*; London, UK: J. Murray, 1859.

[49] Mendel, G. *Experiments in Plant Hybridization* (pp.1-39). Brünn: Proceedings of the Natural History Society of Brünn. 1865, 3-47.