Extracting Three-Dimensional Cellular Automaton for Cement Microstructure Development using Gene Expression Programming

Zhifeng Liang, Bo Yang*, Lin Wang, Xiaoqian Zhang, and Nana He
Shandong Provincial Key Laboratory of Network based Intelligent Computing, University of Jinan,
Jinan, 250022, China
*Corresponding author
E-mail: yangbo@ujn.edu.cn

Ajith Abraham
[1]IT4Innovations, VSB-Technical University of Ostrava,
Ostrava, Czech Republic
[2]Machine Intelligence Research Labs (MIR Labs),
WA, USA
Email: ajith.abraham@ieee.org

***Abstract*—A three-dimensional CA model for the simulation of Portland cement microstructure development has been developed in this paper. The Gene Expression Programming (GEP) algorithm is employed as the learning algorithm to evolve the transition rule reversely from the microstructure development characteristic data due to hydration reactions. The characteristic data is extracted from 8-bit gray images that based on the processing of real cement acquired by Micro Computed Tomography (micro-CT) technology. Starting with initial micro-CT image, cement microstructure evolution images of 28 days is constructed through CA rule discovered by GEP. The experimental results show that this model with the CA rule designed by GEP has higher agreement between the model predictions and experimental measurements for degree of hydration than other models. Furthermore, this model still has good generalization ability when changing the water-cement ratio and chemical composition.**

Keywords-three-dimensional cellular automaton; cement; microstructure development; gene expression programming; reverse modeling

I. INTRODUCTION

The microstructure development affects the cement structure and strength directly. A better understanding of the microstructure development evolution process occurring at each day is necessary to increase the predictability of the cement performance. Meanwhile, the study of hydration process offers scientific challenges in multi-scale theoretical modeling methods. Therefore, it is very significant to study the process of cement microstructure development both in academic and practical field [1].

Initial work on the study of microstructure of composite structure based on computer modeling and simulation published in 1985 by Wittmann [2]. Then, in 1986, Jennings proposed the microstructural model attempted to use a spheres to expressed microstructural evolution process which marking the start of a new era of cement modeling [3]. Based on this model, researchers from Delft University of Technology start a further study in 1987 which paved the way for the production of following powerful model called HYMOSTRUC [4]. This model is developed by Breugel in 1995 which joins the hydration kinetic into the microstructural evolution process considering that the formation of hydration products can affect the hydration rate.

Meanwhile, a group from National Institute of Standards and Technology (NIST) led by Bentz developed a completely different method which combined digital image technology with Cellular Automaton (CA) [5], [6]. CEMHYD3D model put forward by Bentz [7] based on digital image technology and CA methods is proving to be quite an efficient way to simulate the cement hydration and microstructure evolution process. However, the parameters of the model are set only according to the initial condition of cement paste and the three-dimensional cement image is reconstructed by two-dimensional SEM/X-ray image which may result in the deviation from the reality in the simulation process of cement hydration. Therefore, the recently developed Micro Computed Tomography (micro-CT) technology is proposed and used to get the real three-dimensional microstructures image data.

Cellular automata are already commonly used for the modeling of chemical reaction [8], [9] recently, but transition rule of the entire model already known is setting manually based on physical-chemical reaction which is lack of information of real hydration data. Manual derivation of Cellular Automaton (CA) rules model like CEMHYD3D for microstructure development is very difficult because of the extreme complexity of Portland cement hydration. CA rules can be reversely extracted automatically from the observed time series data using evolutionary computation method, like Gene Expression Programming (GEP), Multi Expression Programming (MEP) etc. The difficulties on the manual derivation of the hydration rules and the progress on evolutionary computation promote us to explore and use the evolutionary computation method to extract CA transition rule automatically from the observed hydration degree data of time series.

In this paper, a new type of reverse modeling method with three-dimensional CA is described and used to simulate the process of cement microstructure evolution. GEP is adopted to search for the transition rules of CA from the observed hydration degree data of time series that obtained

978-1-4799-5937-2/14/$31.00 ©2014 IEEE

by micro-CT technology and the experiment get better results than traditional methods.

The rest of the paper is organized as follows: Section II giving an overview of model description after a brief introduction of CA and GEP. Section III introduces the process of data acquisition and feature selection from image data of cement hydration. The results and a discussion on the experiment are offered in section IV. Section V concludes the discussion.

II. REVIEW OF BASIC MODEL

In this section, a brief introduction to the basic knowledge of CA and GEP will be given firstly. Details on modeling framework are discussed later.

A. Cellular Automaton

CA is a discrete model widespread applied to the field of computability theory, mathematics, physics, complexity science, theoretical biology and microstructure modeling. Typically, CA can be regarded as a composition of cell space and transition rules defined on the space. Fig. 1 shows an example of a simple two-dimensional CA. The transition rule is the core of the CA which maps the state of the initial data on a discrete lattice to a future state.

The research of CA currently facing two kinds of problems: one is given specific rules and initial configuration to get the evolution results. The other one is giving the evolution results of series time data to get the specific rules reversely. It is very easy to obtain the next state of CA when the rule and the initial configuration are appointed. However, the inverse problem of identifying the rule from a given data is quite hard. The Fig. 2 shows that. Mitchell et al. use genetic algorithm as a tool to evolving CA rules to perform computations: mechanisms and impediments firstly [10]. Koza et al. then studied on the generation of simple random number use genetic programming and CA [11]. Ferreira used GEP to extract cellular automata rules for the density-classification problem [12].

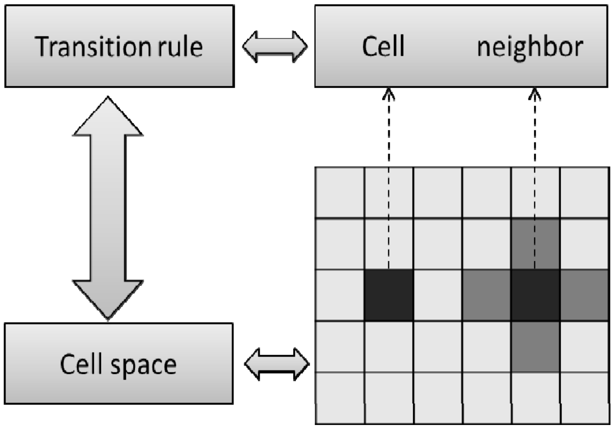

Figure.1 Simple composition of CA

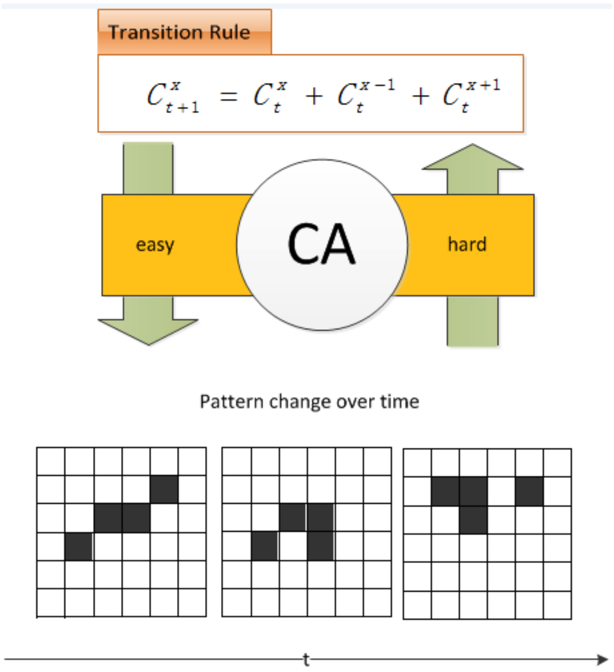

Figure.2 The transformation between data and rules

B. Gene Expression Programming

GEP is a member of Genetic Programming (GP) family that was developed by Ferreira [13] searching for the computer program that solve the problem. It is an important evolutionary technique intended for solving computationally difficult problems like symbolic regression, classification, etc. It works with two entities: chromosomes and expression trees. Candidate solution is represented by an expression tree which is similar with GP tree.

Consider, for example, the expression:

$$\sqrt{(a-b)(c+d)} \quad (1)$$

It can be represented as an expression tree just like following Fig. 3:

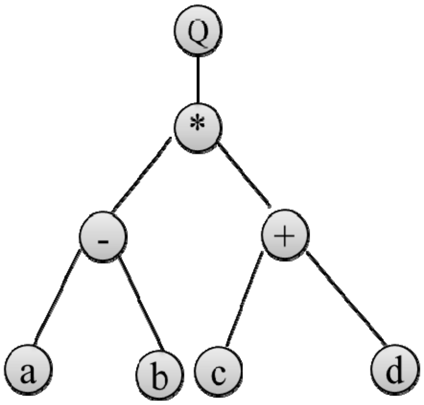

Figure.3 expression tree of GEP

C. Modeling Framework

This paper put forward a real three-dimensional microstructure model for cement microstructure development process. This model consists of the following parts: data acquisition of cement hydration which will be included in section III. The other part is the construct of the CA model. We can see the entire process in Fig. 4.

The data acquisition stage provided training and testing data set for the CA model. For each voxel in three-dimensional microstructure images of a certain time points, extracting the gray feature value of the voxel and the neighborhood voxel value as part of input parameter. Through certain rules or function, the gray feature value is mapped to the value of voxel in the next time points. CA has ability to produce surprisingly complex behavior from simple rules. When the models have been obtained, given an initial cement image, evolution process of the cement microstructure in each day can be simulated successfully.

In our CA model, 5-tuple model is used which is expressed as $A = (L, d, S, N, f)$. L and d constitute a cell space of CA. The cellular automaton we address in this paper consists of a three-dimensional grid of cells, each in one of 256 states mapped to the gray value. The cells which represented with S live on the grid. Each cell has a state and a neighborhood N. In this paper, the neighborhood has a 26 neighbor in cube space. The rule f is just what we want get from the experiment data.

From the above, for our cement models, each cell represents a small pixel of cement microstructure. The values of the variables at each lattice are updated by rule which define by f discovered by GEP. The rule is applied to each cell in one time step and generates the new state of the system at the next step.

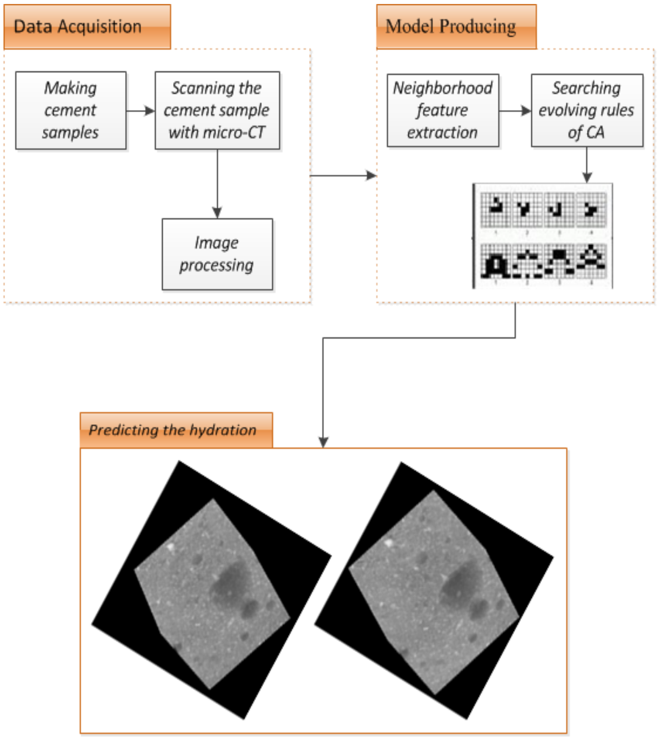

Figure.4 Frame of 3D microstructure modeling for cement hydration

The CA rules which developed by D.P. Bentz in the model of CEMHYD3D is determined by the hydration reactions that occur between cement and water. Using a series of chemical equations as the mapping function, CEMHYD3D can simulate the process of the cement hydration. But for the data which get from micro-CT, the real substance of each voxel is uncertain, therefore, the function cannot get from the chemical equations. For this, GEP algorithm is used to extract function from the real microstructure image data of the cement hydration.

For any CA based model, rules must be selected to define the microstructure change of state occurring at each day. This requires a detailed understanding of the chemical reactions. In our model, chemical reactions are no longer being considered any more. This model is a complete computational intelligence simulation model.

III. DATA ACQUISITION

Methods are presented for generating training and testing dataset for GEP in this section. These are derived from micro-CT technology.

Three types of Portland cement specimen which identified by A, B, and C were used in our experiment with different chemical composition. Chemical compositions are listed in Table I with which weight ratios of C_3S, C_2S, C_3A and C_4AF are calculated below in Table II. Specimens cement labeled by $A_{0.35}$, $A_{0.45}$ and $B_{0.35}$ are constructed as training dataset and the rest of cement sample labeled $C_{0.45}$ is used as testing dataset. The dataset in different subscript is set up in order to compare to the influence of different water-to-cement ratio (w/c) which means that 0.35 and 0.45 is water-to-cement ratio.

First of all, we obtained a large number microstructure images at different time in process of cement hydration according to micro-CT technology. Generally, the hydration data of 28 days is chosen as a reference to study the cement microstructure evolution process, because the degree of cement hydration after 28 days maintain at a low level and the cement performance have been determined basically. Therefore, the images used in our experiment is chosen at following days: 1D(day), 2D, 3D, 4D, 5D, 6D, 7D, 14D, 21D, and 28D.

After a series of image processing like image enhancement and three-Dimensional image registration, we get three-dimensional cement hydration images in each day mentioned above with size of 140*140*140.

Then we developed software to extract the dataset from these images. The dataset is the gray feature value of the voxel and 26 neighborhood voxel value. In addition, we added three feature selected by our experience that can be found in Table III.

As shown in Table III, p_i is the 26 neighborhood-valued of the current pixel values. μ is mean value of neighborhood Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Also, t_i is the ratio between p_i and the sum of 26 neighborhood-valued.

TABLE I. CHEMICAL COMPOSITION OF DIFFERENT CEMENT

Composition	**Specimen**		
	A	***B***	***C***
CaO (%)	63.89	63.74	64.74
SiO_2 (%)	22.58	20.93	20.72
Fe_2O_3 (%)	3.03	3.77	3.74
Al_2O_3 (%)	4.67	4.45	4.66
MgO (%)	2.46	2.51	3.34
SO_3 (%)	1.75	1.56	1.38
K_2O (%)	1.2	0.93	0.54
Na_2O (%)	0.17	0.23	0.19

TABLE II. WEIGHT RATIOS OF C3S, C2S, C3A,C4AF

Composition	**Specimen**		
	A	***B***	***C***
C_3S (%)	45.12	56.50	60.34
C_2S (%)	30.79	17.47	13.97
C_3A (%)	7.25	5.42	6.03
C_4AF (%)	9.21	11.46	11.37

TABLE III. NEIGHBORHOOD FEATURES FOR CA MODEL

Symbol	**Equation**	
	Meanings	***Description***
μ	mean value of neighborhood	$(\frac{1}{27}\sum_{i=0}^{26} p_i)$
σ	variance value of neighborhood	$\sqrt{\frac{1}{27}\sum_{i=0}^{26}(p_i-\mu)^2}$
e	entropy value of neighborhood	$\sum_{i=0}^{26} t_i(1-t_i)$

IV. RESULTS AND DISCUSSION

This section presents results of simulations made on the CA model with the transition rule derived by GEP. GEP was able to find a better rule for the given data compared to the result getting through Line Regression, Multilayer Perceptron (MLP) and SMOreg model.

The training dataset used in our experiment which have been obtained from the last section is consists of two kinds of cement with different composition and w/c.

TABLE IV. GEP PARAMETERS USED IN THIS PAPER

Parameter	**Values**
Crossover	One point crossover
chromosome number	100
Number of Genes	4
linking function	+
Generation	10000
Head size	10
operators	+,-,*,/,ln,exp,pow

TABLE V. EXPERIMENTAL RESULTS FOR TESING DATASETS

Model	**RMSE**
Line Regression	20.87
Multilayer Perceptron	29.01
SMOreg	20.72
GEP	19.74

The parameters that used in GEP are given by the Table IV. The training sample size is set to 6400 while testing sample size is 205.

When the equation of evolution has finished, the best rule *y* can be obtained after the decoding and reduction of the best chromosomes in the history.

$$y_1 = \frac{c_1 x_2}{{x_4}^2 e^{\sqrt{\log x_{26}}}} \tag{2}$$

$$y_2 = \sqrt[4]{\frac{c_3 + x_{24} + x_{17} - x_4}{\frac{x_{19}}{x_7} - \sqrt{c_2}}} \tag{3}$$

$$y_3 = \frac{\sqrt[4]{x_{16}(\mathrm{x}_4 - \mathrm{x}_{26} - \mathrm{c}_5)}}{c_4} \tag{4}$$

$$y_4 = \sqrt[4]{(\mathrm{c}_6 - \mathrm{x}_1)(\mathrm{c}_7 - \mathrm{x}_{23}\mathrm{c}_8(\mathrm{x}_{25} - \mathrm{c}_9))} \tag{5}$$

$$y = y_1 + y_2 + y_3 + y_4 \tag{6}$$

In the equations above, c_1, c_2 ... c_9 are coefficients with the values of 7.34, 6.32, 13.04, 5.71, -0.024, 9.94, -20.63, 1.67, 14.75, respectively.

In order to evaluate the performance of the model, we use root mean square error (RMSE) to evaluate the fitness of each individual which can be found in Table V. The result of

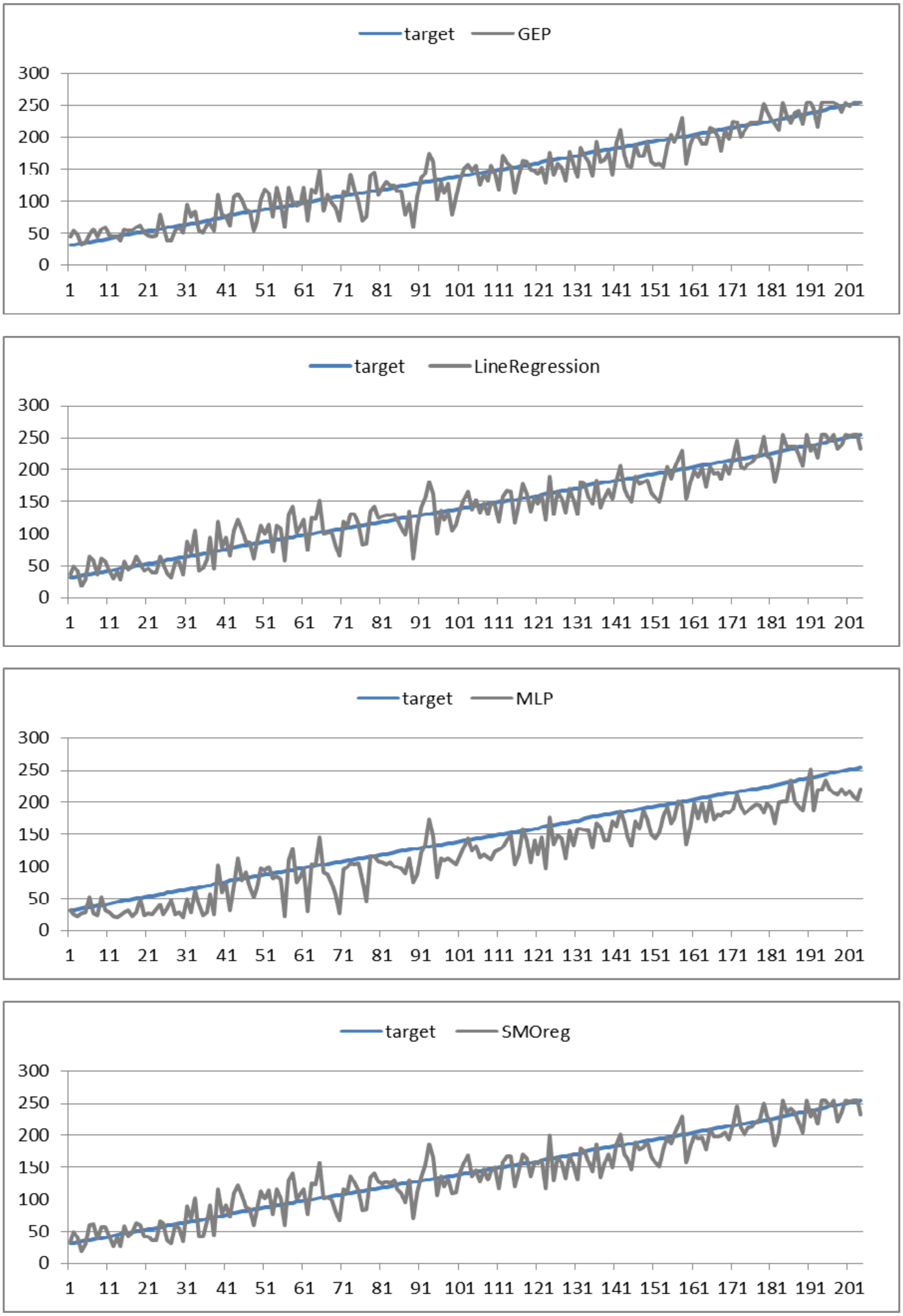

Figure.5 the target and the prediction result

Line Regression, MLP and SMOreg model is get from WEKA.

Then the output and the prediction error with 205 testing data are shown in Fig. 5. The abscissa of Fig. 5 is the numbering of testing data while the ordinate is the output and prediction value.

From Fig. 5, we can see that GEP algorithm can get a more accurate model for the process of microstructure evolution than other model. Compared to the CEMHYD3D, this model obtains more realistic results which can reflect the real process of cement hydration. For all that, there is still one thing should be concerned, GEP algorithm needs lots of running time to get the model.

V. CONCLUSION

In this paper, the method for modeling the cement microstructure development inversion problem based on CA with the algorithm of GEP to search the transition rule was discussed. The experimental result that we obtained can not only fit the training data， but also have a good generalization ability.

We plan to further improve our method to make detection more accurate and robust, by using Multi Expression Programming (MEP). Also, Particle Swarm Optimization (PSO) algorithm is used to find the combinations of the coefficient in the future.

ACKNOWLEDGMENT

This work was supported by National Key Technology Research and Development Program of the Ministry of Science and Technology under Grant 2012BAF12B07-3. National Natural Science Foundation of China under Grant No. 61173078, No. 61203105, No. 61373054, No. 61173079, No. 61070130, No. 81301298. Shandong Provincial Natural Science Foundation, China, under Grant No. ZR2010FM047, No. ZR2012FQ016, No. ZR2012FM010, No. ZR2011FZ001,No. ZR2010FQ028. Jinan Youth Science &Technology Star Project under Grant No. 2013012.

REFERENCES

[1] J. W. Bullard, H. M. Jennings, R. A. Livingston, A. Nonat, G. W. Scherer, J. S. Schweitzer, K. L. Scrivener, and J. J. Thomas, “Mechanisms of cement hydration,” *Cem. Concr. Res.*, vol. 41, no. 12, pp. 1208–1223, Dec. 2011.

[2] F. H. Wittmann, P. E. Roelfstra, and H. Sadouki, “Simulation and analysis of composite structures,” *Mater. Sci. Eng.*, vol. 68, no. 2, pp. 239–248, Jan. 1985.

[3] H. M. Jennings and S. K. Johnson, “Simulation of Microstructure Development During the Hydration of a Cement Compound,” *J. Am. Ceram. Soc.*, vol. 69, no. 11, pp. 790–795, Nov. 1986.

[4] K. van Breugel, “Numerical simulation of hydration and microstructural development in hardening cement-based materials (I) theory,” *Cem. Concr. Res.*, vol. 25, no. 2, pp. 319–331, Feb. 1995.

[5] D. P. Bentz, P. V. Coveney, E. J. Garboczi, M. F. Kleyn, and P. E. Stutzman, “Cellular automaton simulations of cement hydration and microstructure development,” *Model. Simul. Mater. Sci. Eng.*, vol. 2, no. 4, p. 783, Jul. 1994.

[6] D. P. Bentz, “Three-Dimensional Computer Simulation of Portland Cement Hydration and Microstructure Development,” *J. Am. Ceram. Soc.*, vol. 80, no. 1, pp. 3–21, Jan. 1997.

[7] D. P. Bentz, “CEMHYD3D: A three-dimensional cement hydration and microstructure development modelling package. Version 2.0,” *Natl. Inst. Stand. Technol. Interag. Rep.*, vol. 7232, 2000.

[8] L. B. Kier, *Cellular automata modeling of chemical systems: a textbook and laboratory manual*. Dordrecht: Springer, 2005.

[9] A. C. J. de Korte and H. J. H. Brouwers, “A cellular automata approach to chemical reactions; 1 Reaction controlled systems,” *Chem. Eng. J.*, vol. 228, pp. 172–178, Jul. 2013.

[10] M. Mitchell, J. P. Crutchfield, and P. T. Hraber, “Evolving cellular automata to perform computations: mechanisms and impediments,” *Phys. Nonlinear Phenom.*, vol. 75, no. 1–3, pp. 361–391, Aug. 1994.

[11] J. R. Koza, “Evolving a Computer Program to Generate Random Numbers Using the Genetic Programming Paradigm.,” in *ICGA*, 1991, pp. 37–44.

[12] C. Ferreira, *Gene expression programming: mathematical modeling by an artificial intelligence*, 2nd rev. and extended ed. Berlin□; New York: Springer-Verlag, 2006.

[13] C. Ferreira, “Gene expression programming: a new adaptive algorithm for solving problems,” *ArXiv Prepr. Cs0102027*, 2001.