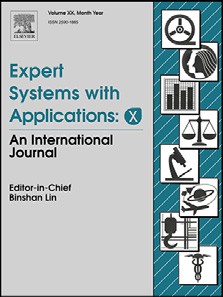
[Expert Systems with Applications: X 2 (2019) 100006](https://doi.org/10.1016/j.eswax.2019.100006)

Contents lists available at [ScienceDirect](http://www.ScienceDirect.com/)

Expert Systems with Applications: X

journal homepage: [www.elsevier.com/locate/eswax](http://www.elsevier.com/locate/eswax)

International roughness index prediction based on multigranularity fuzzy time series and particle swarm optimization

Wei Li[a,](#_bookmark0) Ju Huyan [b,](#_bookmark1)[∗,](#_bookmark2) Liyang Xiao[a,](#_bookmark0)[∗,](#_bookmark2) Susan Tighe [b,](#_bookmark1) Lili Pei [a](#_bookmark0)

a *School of Information Engineering, Chang’An University, Xi’an, Shaanxi 710064, People’s Republic of China*

b *Centre for Pavement and Transportation Technology (CPATT), Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1*

a r t i c l e i n f o a b s t r a c t

*Article history:*

Received 29 January 2019

Revised 22 May 2019

Accepted 23 May 2019

Available online 24 May 2019

*Keywords:*

International roughness index Fuzzy time series Multigranularity

Automatic clustering Particle swarm optimization

The effective prediction of pavement performance trends can help in achieving the cost-effective man- agement of pavements over their service life. The international roughness index (IRI) is a widely used pavement performance index, which can be considered as a time-dependent variable in terms of scien- tific modeling. This research aims to develop an innovative IRI prediction model based on fuzzy-trend time-series forecasting and particle swarm optimization (PSO) techniques. Raw datasets extracted from the Long-Term Pavement Performance database are used for model training, testing, and performance assessment. First, IRI values are divided into different granular spaces, which are considered as the prin- cipal factor and subfactors. In addition, the multifactor interval division method is proposed according to the principle of the automatic clustering technique. Next, a second-order fuzzy-trend model and fuzzy- trend relationship classification method are proposed to predict the fuzzy-trend of each factor. Then, the fuzzy-trend states for multiple granular spaces are generated while giving full consideration to various uncertainties. Finally, the PSO technique is used to optimize the performance model while carrying out future IRI forecasting. Comparative experiments are performed using more than 20,000 data items from different regions to verify the effectiveness of the proposed method. The experimental results indicate that the proposed method outperforms other approaches including the polynomial fitting, autoregres- sive integrated moving average, and backpropagation neural network methods in terms of the root mean squared error (0.191) and relative error (6.37%).

© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

### Introduction

In transportation engineering, the topmost consideration should be safety. Therefore, professional scientific effort should be made by both private and public sectors to comprehensively interpret service conditions and the changing mechanism of road roughness levels. The international roughness index (IRI) was initially pro- posed in a research project conducted by the University of Michi- gan ([Cebon, 1993](#_bookmark54)) for monitoring the overall roughness condition of certain pavements. The IRI mainly measures the longitudinal profile condition of a travelled roadway according to the vehicle vibration condition. The most widely used units for measuring the

∗ Corresponding authors.

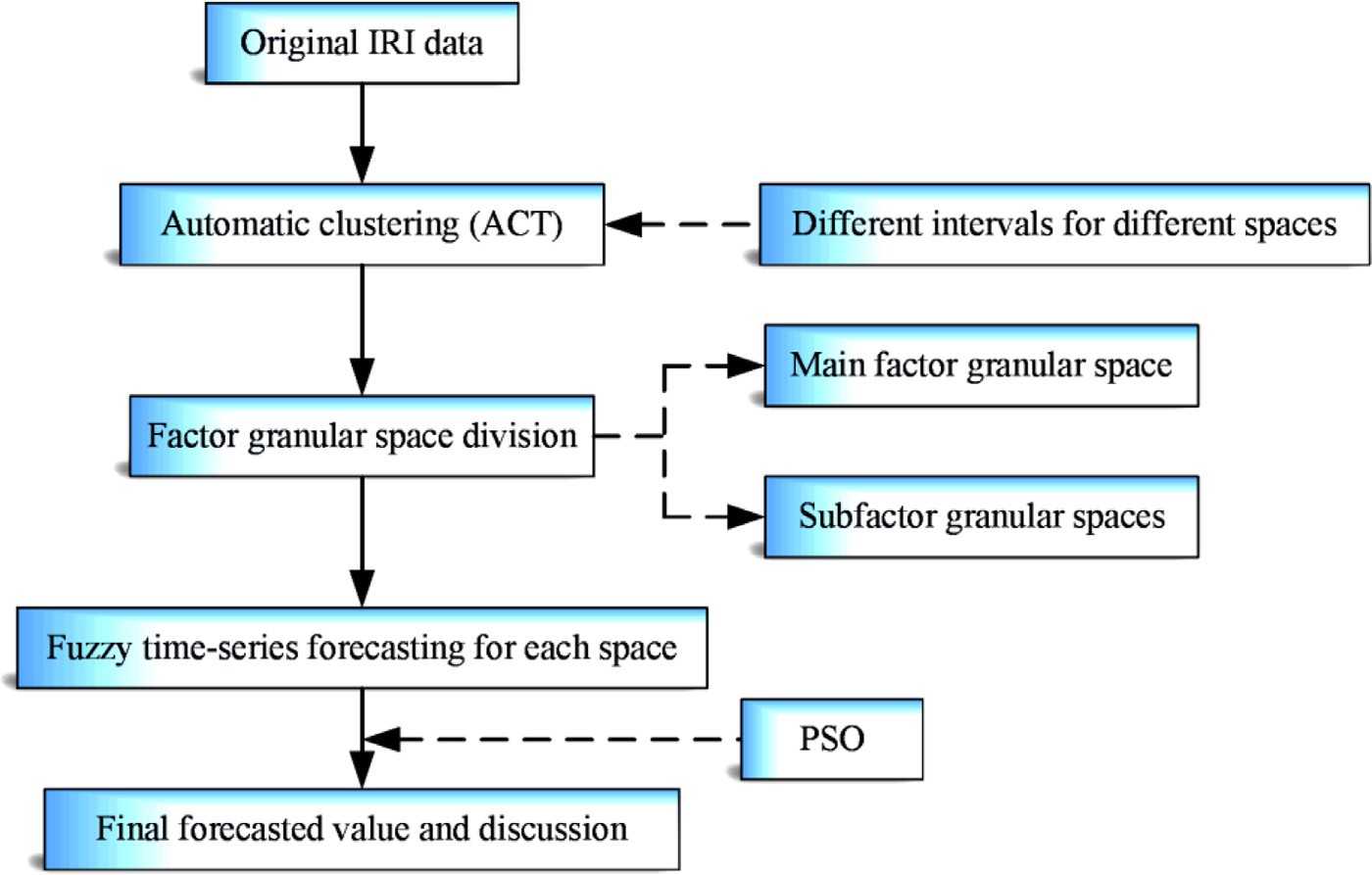
*E-mail addresses:* [grandy@chd.edu.cn](mailto:grandy@chd.edu.cn) (W. Li), [jhuyan@uwaterloo.ca](mailto:jhuyan@uwaterloo.ca) (J. Huyan), [liyang\_xiao@163.com](mailto:liyang_xiao@163.com) (L. Xiao), [sltighe@uwaterloo.ca](mailto:sltighe@uwaterloo.ca) (S. Tighe), [2543487651@qq.com](mailto:2543487651@qq.com) (L. Pei).

[IRI are meter per kilometer and inch per mile (Zhou, Wang, & Lu, 2008).](#_bookmark59)

Over the past few decades, scholars across the world have made significant effort to study the changing IRI mechanisms, which have been then used to analyze pavement performance [deterioration trends (Kargah-Ostadi, Stoffels, & Tabatabaee, 2010; Lu & Tolliver, 2012; Peshkin, Hoerner, & Zimmerman, 2004; Shi- razi, Ayres, Speir, Song, & Hall, 2010; Zimmerman & Peshkin, 2003).](#_bookmark43) [R](#_bookmark47)[ahim, Fiegel, Ghuzlan, and Khumann (2009) analyzed](#_bookmark43) the IRI characteristics of cracked and seated concrete pavements and thereby provided maintenance and rehabilitation guidance for pavement management. However, they did not consider the cli- mate conditions, which have significant influence on the rough- ness progression, especially for hot-mix asphalt overlay pavements ([Perera & Kohn, 2001](#_bookmark44)). [Albuquerque and Núñez (2011)](#_bookmark53) investigated the roughness condition of low-volume roads using multiple re- gression analysis. [Zhao and Guo (2013)](#_bookmark56) predicted the IRI over time using Markov analysis, thus providing a probabilistic IRI prediction

<https://doi.org/10.1016/j.eswax.2019.100006>

2590-1885/© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)



**Fig. 1.** Methodology framework.

[method instead of a deterministic one. Qian, Jin, Zhang, Ling, and Sun (2018) proposed an exponential IRI regression model by mod-](#_bookmark46) eling two parameters *α* and *β* and used the model to predict IRI trends over time. However, the effectiveness and robustness of the abovementioned approaches are not convincing because most re- gression analysis methods assume that the dataset follows certain distributions.

Typical IRI prediction models have been developed by the transportation departments of Mississippi and Washington as part of the Long-Term Pavement Performance (LTPP) program and by other state agencies based on the Mechanistic Empirical Pave- [ment Design Guide (George, 2000; Jackson, Puccinelli, & Engineers, 2006; Oﬃcials, 2008). The development of these models is a re-](#_bookmark37) markable achievement, and they are still being used for long-term pavement design and management. However, because these mod- els are based on data-driven approaches, they still have the in- [trinsic drawbacks of statistic-based methods (Contreras, Espinola, Nogales, & Conejo, 2003). Thus, the performance of these models](#_bookmark38) relies significantly on the quality and characteristics of the data being used. In addition, traditional approaches are influenced by the low accuracy of data acquisition, and hence, they usually fail to handle the ambiguity and uncertainty inherent in raw datasets ([Reuter, 2007](#_bookmark48)).

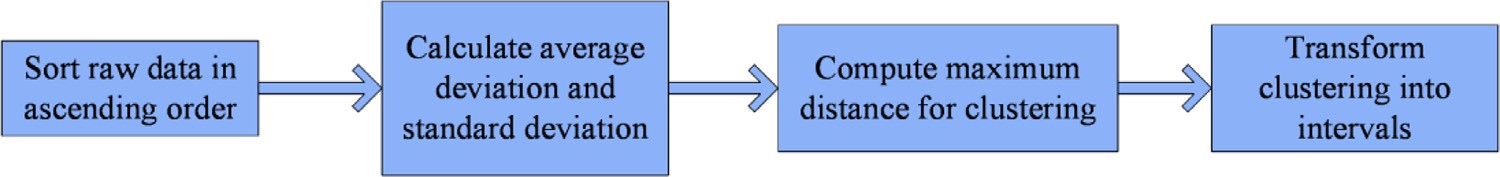
With the advancement of artificial intelligence (AI) techniques, several engineering fields are focusing on using AI approaches to facilitate research and practice tasks. In the field of pave- [ment management, the initial attempt was made by Kargah- Ostadi et al. (2010), who proposed an intelligent method based](#_bookmark43) on artificial neural networks (ANNs) to analyze the deterioration trend of pavements through IRI prediction modeling. Thereafter, researchers have used not only ANNs but also other AI tech- [niques such as support vector machines (Abdelaziz, Abd El-Hakim, El-Badawy, & Afify, 2018; Georgiou, Plati, & Loizos, 2018) to in-](#_bookmark51) vestigate pavement roughness progression. However, these meth- ods consider the IRI as a dependent indicator that is influenced by several other pavement conditions, traﬃc loading, and envi- ronmental factors at a fixed observation time, thereby ignoring the impact of time on the IRI. Time is an invisible space vari- able that can have an abstract existence and can contain a com- prehensive combination of all other visible, traceable, and con- crete influences. Hence, this variable has a direct influence on the deterioration of pavement performance, and hence, it affects

the IRI. Therefore, fuzzy time-series methods should be investi- gated to bridge the gaps in current literature with regard to intel- ligent IRI predictions. Fuzzy time-series prediction methods have been successfully applied to several fields such as stock predic- tion and weather prediction, and these applications serve as in- valuable reference for the IRI prediction research conducted in this study.

Considering the background described above, the present study aims to develop an IRI prediction approach based on an innovative fuzzy time-series analysis. To fulfill this purpose, multifactor and multigranularity analyses are considered for time-series forecast- ing. First, the automatic clustering technique (ACT) is adopted to generate a series interval within each granular space, which should be defined in advance. Next, a second-order fuzzy-trend matrix (SFTM) and fuzzy-trend relationship classification (FTRC) method are proposed to predict the fuzzy-trend of each factor. Then, the fuzzy-trend states of multiple granular spaces are generated while giving full consideration to various uncertainties. Finally, the par- ticle swarm optimization (PSO) technique is employed to forecast the optimal IRI value. Comparative experiments are performed to evaluate the effectiveness of the proposed method.

### Methodology

After a road is opened to traﬃc, its roughness condition tends to vary over the service time. To outline the changing mechanism of road roughness, IRI values are regarded as fuzzy time-dependent variables. [Fig. 1](#_bookmark3) illustrates the framework of the proposed fuzzy time-series clustering IRI prediction methodology. As shown in the figure, the IRI values are first divided into three granular spaces. Then, the data within each granular space are assigned to a se- ries of intervals, with each interval having a membership degree defined by fuzzy theory. Specifically, the value of the membership degree indicates the belonging probability of the concrete IRI value to the corresponding interval. In this research, the initial intervals are obtained through an automatic clustering algorithm. Thereafter, the fuzzy sets are assigned to the divided intervals according to the subordinate relationships of these sets along with the grouping statistics. Finally, the PSO method is utilized to obtain the optimal weight vector and corresponding predicted IRI value. To provide a more interpretative understanding of the methodology, the key fundamental theories are introduced in this section.



**Fig. 2.** ACT processes.

* 1. *Fuzzy time series*

*de*

=

*v* , Σ*n (d*

— *d* − *agv )*2

*di f f*

(5)

[fuzzy time-series analysis. On the basis of this theory, Song and Chissom (1993) developed the first fuzzy time-series model in](#_bookmark50) 1993, which was a remarkable breakthrough. Later, several stud- [ies were conducted on the optimization (Chen & Jian, 2017; Chen, Manalu, Pan, & Liu, 2013; Yu, 2005) and application of](#_bookmark58) time-series analysis. A few representative examples are listed here: [Chen (1996)](#_bookmark57) aimed to optimize the complexity of a fuzzy time-series model, [Huarng (2001)](#_bookmark40) proposed a heuristic algorithm ([Kuo et al., 2009](#_bookmark45)) to improve the accuracy of the prediction result of a fuzzy time series, and [Zadeh (1996)](#_bookmark55) proposed a time-series model optimized for computing with words prediction analysis.

[In 1965, Zadeh proposed the fuzzy set theory (Sets & Zadeh,](#_bookmark49)

[1965; Yao, 2000), which was the theoretical foundation for](#_bookmark49)

*i*=1 *i*+1 *i* diff

*n* − 1 − 1

Here, a fuzzy time series is defined as follows:

Define *U* = {*u*1 *, u*2 *, . . . , un*} as a universe of discourse. Then, a fuzzy set *A* of *U* can be represented as follows:

where *n* denotes the total number of data points. *di (i* = 1*,* 2*, . . . , n)*

denotes the sorted data values with *d*1 *< d*2 *< . . . < di < . . . < dn*.

Step 2: Use [Eq. (6)](#_bookmark6) to calculate the maximum distance:

max \_*distance* = *c* × *devdi f f* (6)

where *c* is a constant weight value. Here, *c* = 0*.*5.

Step 3: Cluster the ascending data following the calculations of

Steps 1 and 2. Let *d*1 be the current cluster and judge whether the next data point belongs to this current cluster or not. If not, assign it to a separate cluster. That is, if *di*+1 − *di* ≤ max\_distance,

then assign *di*+1 to the current cluster. Otherwise, create a new

cluster for *di*+1. Repeat the operations until all the data points are assigned to specific clusters. The clustered results can be expressed

as follows:

*A* =  *fA (u*1 *)* +  *fA (u*2 *)* + *. . .* +  *fA (u*n *)*

*u*1

*u*2

*u*n

(1)

{*d*11*, . . . d*1*a*}*,* {*d*21*, . . . d*2*b*}*, . . .* {*di*1*, . . . din*}*, . . .* ©*dj*1*, . . . djm*}

where *fA* denotes the membership function of the fuzzy set *A*, which is mostly written in the form *fA* : *U* → [0, 1]. *fA* (*ui* ) is the membership degree of *ui* within the fuzzy set *A*, 1 ≤ *i* ≤ *n*.

Assume that a group of fuzzy sets *fi (t)(i* = 1*,* 2*, . . .)* are de- fined in the universe of discourse Y*(*t*)(t* = *. . . ,* 0*,* 1*,* 2*, . . .)*. *F*(*t*) rep- resents the collection of *fi (t)(i* = 1*,* 2*, . . .)*. Then, the collection F*(*t*)* = { *f*1 *(t ), f*2 *(t ), . . .*} is called a fuzzy series set in the discourse *Y (t)(t* = *. . . ,* 0*,* 1*,* 2*, . . .)*.

Based on the above definition, the *n*th−order fuzzy time-series relationships for the fuzzy sets F*(t* − 1*),* F*(t* − 2*), . . . ,* F*(t* − *n)* can be expressed as

F*(t* − *n),* F*(t* − *n* + 1*), . . . ,* F*(t* − 1*)* → F*(*t*)* (2)

Therefore, given F*(*t*)* = *A j , F (t* − 1*)* = *Ai*1*, . . . , F (t* − *n)* = *Ain* , the following logical relationship can be obtained:

*Ain, . . . , Ai*2*, Ai*1 → *A j* (3)

where *Ain, . . . , Ai*2 *, Ai*1 are called the current states, which should be known values in a certain problem. *Aj* is called the next state, which is to be predicted by the developed model.

Step 4: Convert the clustering results to continuous intervals.

[Fig. 3](#_bookmark8) presents the process of converting clusters to intervals. The detailed steps are as follows. First, convert the first cluster

{*d*11 *, ... , d*1*a* } to an interval [*d*11 , *d*1*a* ). Then, make the following

two judgements: (1) Judge whether the current interval is [*di, dj* ) with the current class of {*dk, dl* }. If *dj* ≥ *dk* , then convert the cur- rent class {*dk, dl* } to intervals [*dj, dl* ). At the same time, use [*dj, dl* ) as the current interval and the next cluster, which is {*dm, dn*},

as the current class. However, if *dj < dk* , then convert the current

class {*dk, dl* } to interval [*dk, dl* ) and create a new interval [*dj, dk* ) between [*di, dj* ) and [*dk, dl* ). Next, use [*dk, dl* ) as the current inter-

val and the next cluster, which is {*dm, dn*}, as the current class. (2) Judge whether the current interval is [*di, dj* ) with the current class of {*dk* }. If so, convert the current class {*di, dj* } to interval [*di* , *dk* ).

At the same time, use [*di* , *dk* ) as the current interval and the next

cluster, which is {*dm, dn*}, as the current class. According to this principle, process all the data points within the dataset and thus generate all the intervals.

Step 5: Calculate the medium values for all the intervals as fol- lows:

## *lowi* + *upperi*

* 1. *Interval division using ACT*

*mid*\_*valuei* =

# 2

(7)

The ACT ([Hao, Wang, & Tuo, 2015](#_bookmark39)) is characterized by its adaptive clustering capability, which can help in minimizing the bias caused by manual parameter setting. Therefore, the ACT is adopted to generate the multiple intervals for different granular spaces. [Fig. 2](#_bookmark4) illustrates the key procedures of the ACT, which are data sorting, distance calculation, automatic clustering, and inter- val generation.

Step 1: Sort the raw data in the ascending order, with equal values appearing only once. Then, calculate the average deviation (*agvdiff*) and standard deviation (*devdiff*) of the sorted dataset using

[Eqs. (4)](#_bookmark7) and [(5)](#_bookmark5).

where *lowi* denotes the lower bound of the *i*th interval. Here, the

superscript *i* denotes the upper bound of the *i*th interval.

* 1. *Granular computing*

Granular computing (GrC) is an umbrella concept and compu- tational paradigm for information processing that simulates hu- man thinking mechanisms. Different formalisms and information processing methodologies are included in GrC methodologies. The hierarchy structure of the computing organization mechanism of granules provides GrC with the ability to interpret complex prob- lems from multiple perspectives. As illustrated in [Fig. 4](#_bookmark9), the entire GrC structure usually includes several layers with crossed informa-

*v* Σ*n*

*ag di f f* =  *i*=1

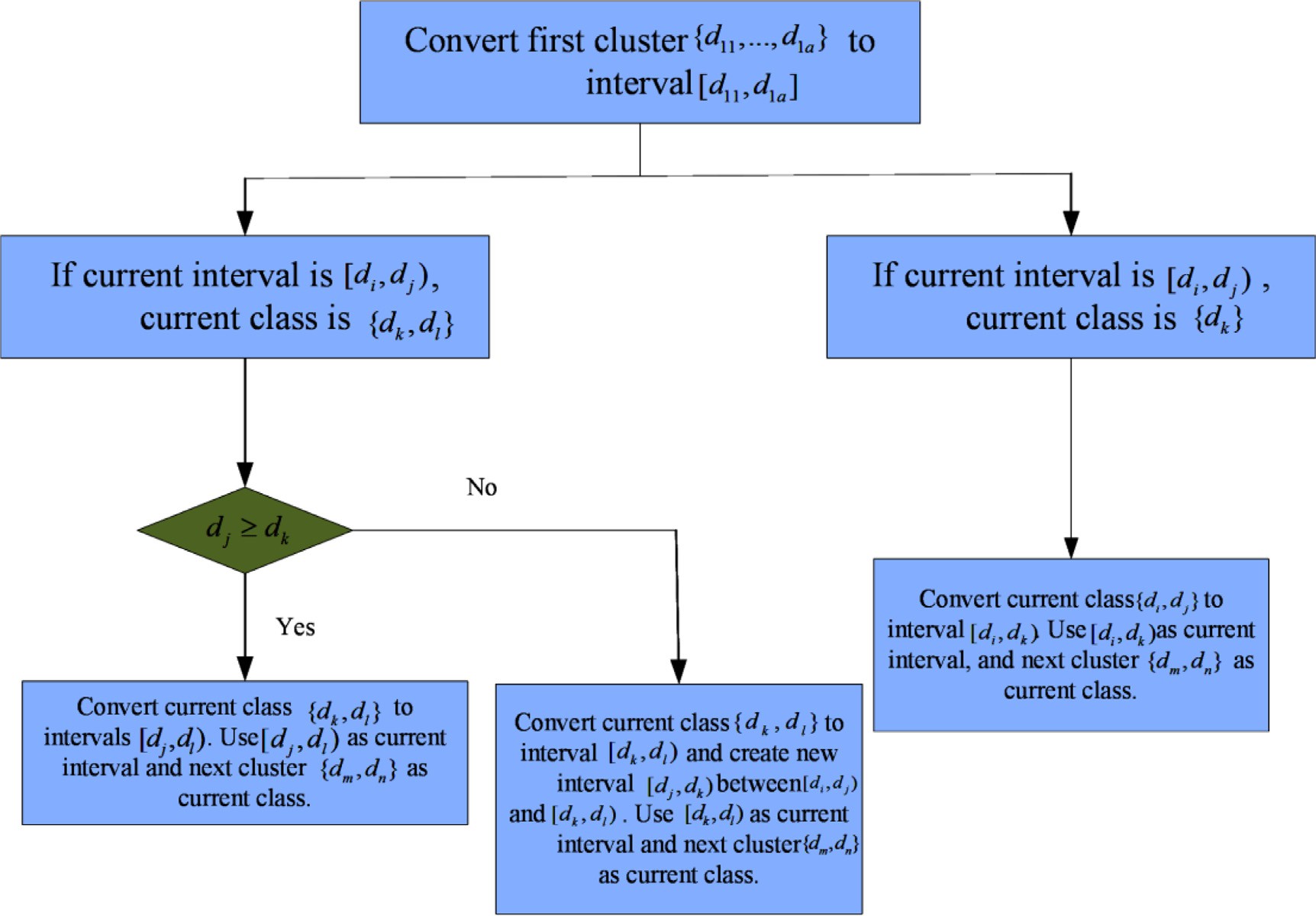
*n* − 1

(4)

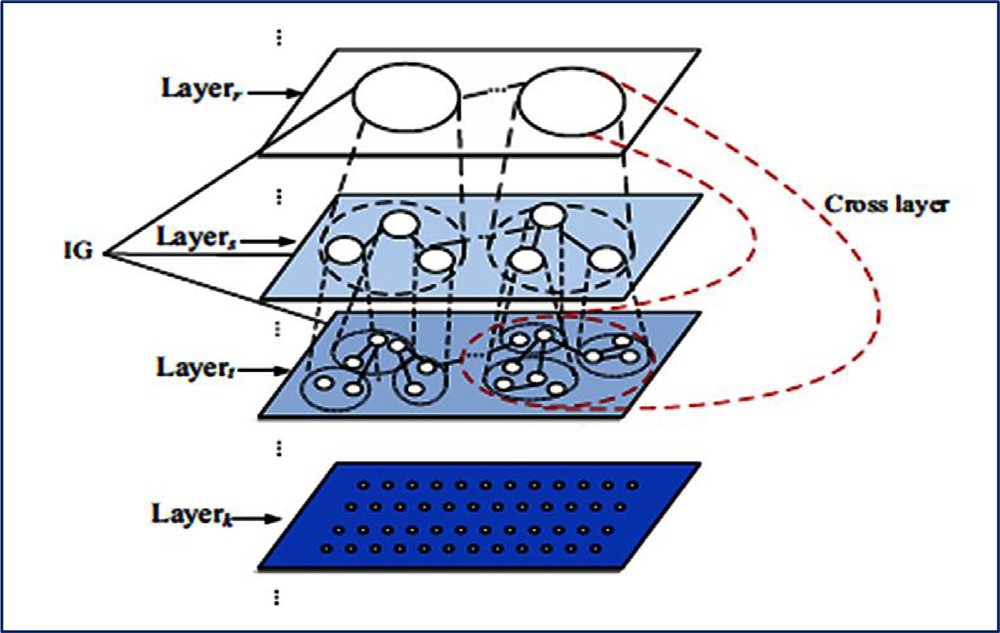
*(di*+1 − *di )*

tion sharing between the granules. In this figure, the deeper the

layers, the finer are the unit data items, which are shown as the



**Fig. 3.** Process of converting clusters to intervals.



**Fig. 4.** Schematic of multilayer granular structure.

dots in the *k*th layer. The hierarchy property of the granular struc- ture can be observed as well, which characterizes the novelty of GrC compared to other methods.

Generally, GrC is realized based on three basic notions: gran- [ules, granular space, and hierarchy structures (Wang, Zhang, Ma, & Yang, 2011; Yao, 2005). Granules are the unit elements of a GrC](#_bookmark52) model and are distributed in their specific granular spaces. The in- dividual granular space (father granular space) can be composed of several subgranular spaces (child granular spaces), and both the father and child granular spaces themselves can be hierarchically distributed as well. Thus, the comprehensive hierarchy of multiple granular spaces can be used to model the multifactor and multi- scale organized datasets, especially for handling problems that in- volve both certainty and uncertainty influences that should be con- sidered simultaneously ([Chen & Tanuwijaya, 2011](#_bookmark60)).

For GrC-based system development, several computing models have been proposed as a type of knowledge interpretation method-

ology. Popular GrC models include the fuzzy set theory, rough set theory, quotient space, and cloud model. Each of them has a spe- cific hierarchy structure suitable for certain problems. The fuzzy set theory is mainly used for handling ambiguous data and is based on the granularity of information and if–then structures. The rough set theory is a mathematical tool for interpreting incomplete data and unveiling the hidden properties of a dataset. Hierarchy rough set structures are a typical characteristic of rough set analysis. The quotient space is characterized by quotient typology, and new space is created for the given data. The cloud model method in- volves forward and backward cloud transformation. Hence, the hi- erarchy structure of the cloud model is generated by this bidirec- tional transformation concept. These four typical GrC models are widely used in literature. Each has its own computing mechanism, but they share the basic ideas of GrC. In this research, the fuzzy set model is selected for IRI prediction analysis because it is more computationally eﬃcient than the other approaches.

Specifically, each IRI measurement factor is considered as an in- dividual granule. Therefore, three granules can be defined by three IRI factors, which are the principal factor and two subfactors. Then, we define three granular spaces: (1) the first granular space con- taining only the principal factor, (2) the second granular space con- taining the principal factor and one of the subfactors, and (3) the third granular space containing the principal factor and both the subfactors.

* 1. *PSO*

PSO ([Hao et al., 2015](#_bookmark39)) is a heuristic model optimization algo- rithm based on the simulation of animal movement behavior. The innovative idea behind this method is that animals are able to se- lect the optimum speeds and directions in moving toward their de- sired destinations. The heuristic aspect of the PSO method is that no assumptions are required for handling the specific tasks when using this method.

The PSO algorithm is a global random search algorithm based on swarm intelligence and hence is more intelligent than tradi- tional methods. The main idea behind this method is the simula- tion of the migration and clustering behavior during bird foraging. Various organisms in nature have certain group behaviors; thus, one of the main research areas of artificial life is to explore the group behaviors of natural organisms, thereby constructing intelli- gent models for solving practical problems with scientific perspec- tives.

In most situations, PSO is used to optimize the solutions of other methods instead of providing independent models. Specifi- cally, potential solutions to a certain problem are regarded as birds, called particles, in the solution space. Each particle has a fitness value that is determined by the defined fitness function and has a velocity that determines the direction and distance from which the particle is to move toward the optimum position. When the spe- cific optimization problem has been established, a group of ran- dom particles (random solutions) is generated to initialize the par- ticles. Then, the moving velocity and positions of the particle are iteratively updated to find the optimal solution. In each iteration, the particle updates itself by tracking two extremums: the optimal solution found by the particle itself, which is called the individual extremum, and the optimal solution found among the entire pop- ulation in the current state, which is called the global extremum. Finally, the optimized model is obtained when the satisfied fitness value is achieved.

Therefore, the primary consideration of utilizing PSO in this re- search is based on the three abovementioned reasons, which can be summarized as follows: (1) PSO is an intelligent algorithm that outperforms traditional methods, yet few studies have used this approach to analyze pavement roughness conditions. (2) PSO is best utilized when it is combined with other methods to solve cer- tain optimization problems instead of being independently used for model building. In this manner, PSO can serve as an effec- tive approach to solve complex optimization problems. (3) The PSO method has been proven to be significantly effective in solving dif- ferent types of problems; hence, this research aimed to determine the effectiveness of the PSO method in solving the problem of IRI prediction.

Specifically, a particle swarm is formed by a group of particles that represent the potential solutions. Each individual particle has its initial position and moving velocity. By calculating the objec- tive function, the particle should update its position based on the velocity toward the best position, which represents the optimum solution of a certain problem. To ensure that the final solution is optimum from a global perspective, the individual local best posi- tion (*Pbest* ) and global best solution (*Pgbest* ) should be updated re-

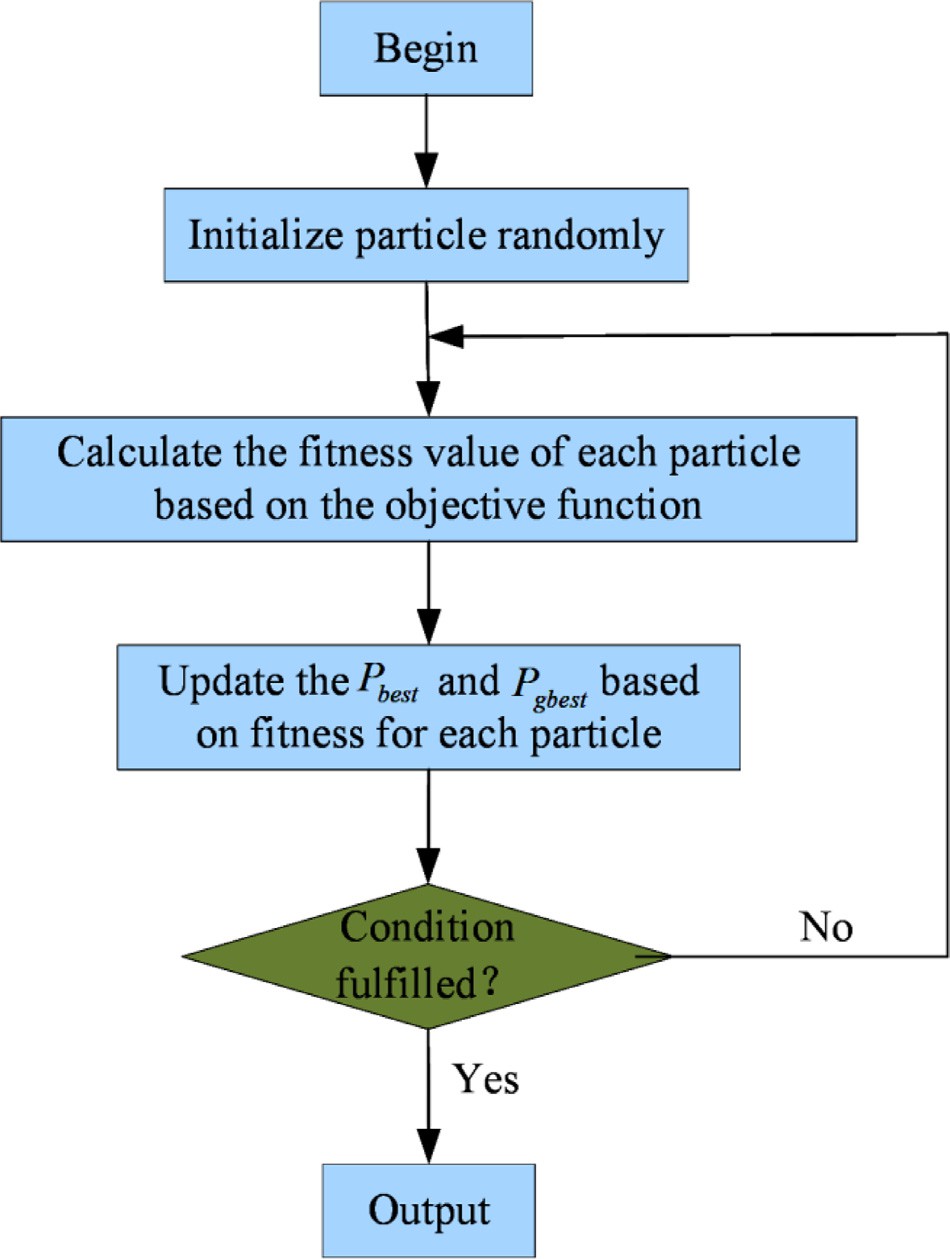
peatedly during the model optimization process. The global best

solution is updated based on the local best solution, which is ob- tained by revising the current position based on the velocity and direction of the particle. The velocity and position update functions for particle *i* are defined in [Eqs. (8)](#_bookmark11) and [(9)](#_bookmark12), respectively.

¡ ¢

*Vi,k* = *w* × *Vi,k*−1 + *C*1 × *r*1 × *Pbest,i* − *Wi,k*−1 + *C*2 × *r*2

×¡*Pgbest* − *Wi,k*−1¢ (8)



**Fig. 5.** PSO flowchart.

### Multigranularity-combination-based IRI prediction

This research proposed a state-of-the-art IRI prediction model based on a fuzzy time-series forecasting methodology facilitated by multigranularity PSO. As shown in [Fig. 6](#_bookmark14), the prediction process is divided into five main steps. In this section, the detailed calcula- tions in each step are discussed.

Step 1: ACT-based interval division: [Fig. 7](#_bookmark15) shows the interval di- vision process based on the ACT. Assume that the universe of dis- course of the principal factor is divided into *p* intervals, which are represented as *u*1 *, u*2 *, . . . up*. If the *j*th subfactor belongs to the in-

terval represented as *mj* , then assign it to this interval. This aﬃlia- tion can be represented as *v j,*1*, v j,*2*, . . . , v j,m* .

*j*

Step 2: FTRC realization: This procedure can be divided into four steps ([Fig. 8](#_bookmark16)): (1) define the fuzzy sets, (2) fuzzify data, (3) construct an SFTM, and (4) build FTRC for each factor. The details of these steps are as follows.

1. Define fuzzy sets for both the principal factor and sub- factors: Consider that *A*1 *, A*2 *, . . . , Ap* represent the fuzzy set of the principal factor and *B j,*1 *, B j,*2 *, . . . , B j,mj* represent the

fuzzy sets of the *jth* subfactors. These two definitions are ex- pressed in [Eqs. (10)](#_bookmark13) and [(11)](#_bookmark19).

*Pi,k* = *Pi,k*−1 + *Vi,k* (9)

1

*A*1 =

*u*

1

0*.*5

+

*u*2

0*.*5

+

*u*3

0*.*5

+

*u*4

0

*. . .* + 

## *uP*

where *Vi, t* and *Pi, k* respectively represent the velocity and position of the *i*th particle in the *k*th iteration. *w* denotes the inertial weight

0*.*5 1

*A* = +

2

*u*1 *u*2

0*.*5 0 0

+ +  *. . .* + 

*u*3 *u*4 *uP*

coeﬃcient. *C*1 and *C*2 are the constant parameters that are decided based on the properties of a specific problem. In the present case,

we set *C*1 = *C*2 = 2. *r*1 and *r*2 are normalization parameters having Gaussian distributions.

0

*A*3 =

*u*

1

0*.*5 1

+ +

*u*2 *u*3

*. . . . . .*

*. . . . . .*

0*.*5

+

*u*4

0

*. . .* + 

## *uP*

The procedures of the PSO optimization algorithm are based on

the above definitions and explanations, as illustrated in [Fig. 5](#_bookmark10).

0

*AP* =

*u*

1

0 0

+  +

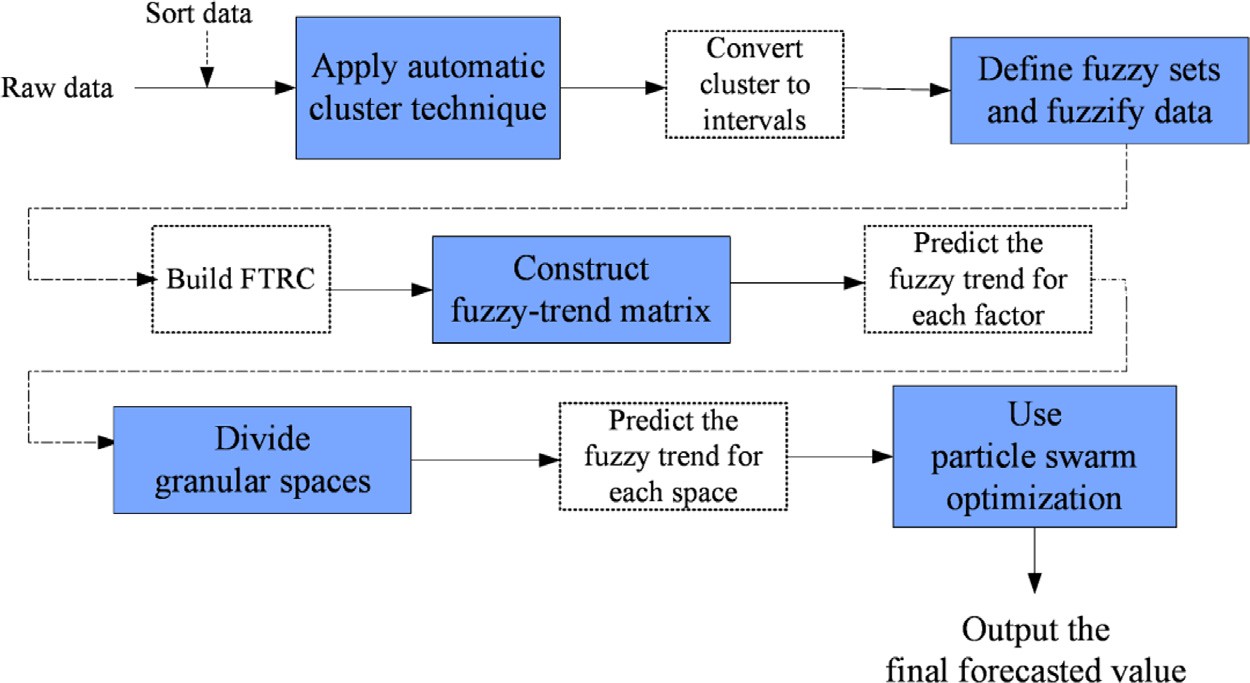
*u*2 *u*3

0 1

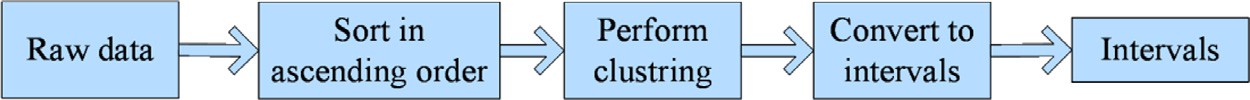
+  *. . .* + 

*u*4 *uP*

# (10)



**Fig. 6.** Illustration of IRI prediction model.



**Fig. 7.** Interval division by ACT.



**Fig. 8.** FTRC realization.

# 1

*Bj,*1 =

*v*

*j,*1

0*.*5

+

*v*

*j,*2

0*.*5

+

*v*

*j,*3

0*.*5

+

*v*

*j,*4

*. . .* + *v*

# 0

*j, mj*

**Table 1**

FTRC realization.

Classification Relationship (*i*2/*k*2 and *i*1/*k*1) Group 1 Descending

# 0*.*5 1

*Bj,* 2 = +

*v v*

*j,*1 *j,*2

0*.*5 0

+ +

*v v*

*j,*3 *j,*4

0

*. . .* +

*v*

*j,mj*

Group 2 Equal

Group 3 Ascending

# 0

*Bj,*3 =

*v*

*j,*1

0*.*5

+

*v*

*j,*2

1

+

*v*

*j,*3

0*.*5

+

*v*

*j,*4

*. . .* + *v*

0

*j,mj*

*. . . . . .*

*. . . . . .*

# 0 0 0 0 1

+

+

+

**Table 2**

Fuzzy-trend matrix.

Group Relationship

*Bj,mj* = *v*

*j,*1

*vj,*2

*vj,*3

*vj,*4

*. . .* + *v*

*j,mj*

# (11)

Descending Equal Ascending

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Sum | Count |  | Sum | Count |  | Sum | Count |  |
| Group 1 | *Sd*(1) | *Nd*(1) |  | *Se*(1) | *Ne*(1) |  | *Sa*(1) | *Na*(1) |  |
| Group 2 | *Sd*(2) | *Nd*(2) |  | *Se*(2) | *Ne*(2) |  | *Sa*(2) | *Na*(2) |  |
| Group 3 | *Sd*(3) | *Nd*(3) |  | *Se*(3) | *Ne*(3) |  | *Sa*(3) | *Na*(3) |  |

where *ui, i* = 1*,* 2*, . . . , p* and *v j,m*

*j*

the ACT.

denotes the intervals divided by

1. Fuzzify data: Based on the fuzzy sets defined previously, as- sign all the recorded historical data to the intervals they be- long to. If a certain value is within the interval of ui, then fuzzify that value into Ai. Similarly, if that value belongs to the interval vj, k, then fuzzify it into Bj, k.
2. Construct a second-order fuzzy-trend matrix (SFTM) for the

fuzzified data: Assume that the fuzzy datasets of the princi- pal factor are *Ai*2, *Ai*1, and *Am* at time points *t* − 2*, t* − 1, and *t*, respectively. Then, the SFTM of the principal factor can be

described as *Ai*2, *Ai*1 → *Am*. Similarly, the SFTM of the subfac- tors are expressed as *Bj, k*2, *Bj, k*1 → *Am*.

1. Build FTRC for each factor: The fuzzy datasets at time points *t* − 2 and *t* − 1 are considered as the current states, whereas that at the time point *t* is regarded as the next state. The

FTRC of the principal factor can be expressed as

(

if i2 *> i*1*(*or k2 *> k*1*), classi f y* SFTM to group 1 *if* i2 = i1 *(* or k2 = k1*), classi f y* SFTM to group 2 *if* i2 *< i*1 *(* or k2 *< k*1*), classi f y* SFTM to group 3

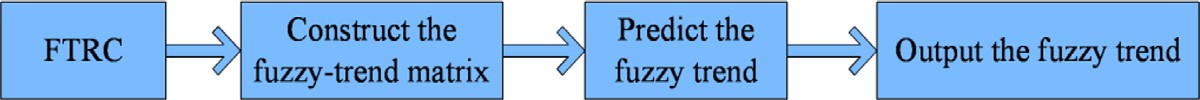
# (12)

Based on this equation, the FTRC can be obtained, as listed in [Table 1](#_bookmark17).

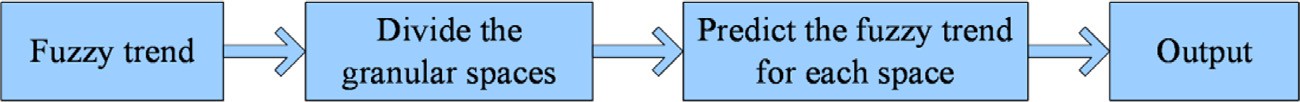
Step 3: Fuzzy-trend calculation for the individual factor: Con- struct the fuzzy-trend matrix ([Table 2](#_bookmark18)), which is used to predict the fuzzy-trend for each factor. [Fig. 9](#_bookmark20) illustrates the procedures of this method.

Based on the fuzzy relationships presented in [Table 1](#_bookmark17), a fuzzy- trend matrix with a size of 3 × 6 is constructed ([Table 2](#_bookmark18)) and as- signed to each factor. In [Table 2](#_bookmark18), *Sd*(*i*) represents the sum of the de- scending trends for group *i*, and *Nd*(*i*) represents the total number of descending relationships. *Se*(*i*) represents the sum of the equal trends of all SFTMs for group *i*, and *Ne*(*i*) represents the total num- ber of equal relationships. *Sa*(*i*) represents the sum of the ascending trends for group *i*, and *Na*(*i*) represents the total number of ascend- ing relationships. The calculation principles are as follows.

Consider the principal factor whose SFTM is expressed as *Ai*2, *Ai*1 → *Am*. If *i*1 *> m*, then update the corresponding values based on



**Fig. 9.** Fuzzy-trend prediction process.



**Fig. 10.** Granular space division process.



**Fig. 11.** Final predicted value calculation.

the following equations:

*V* = *( Sd(i) , Se(i) , Su(i) )* and (2) the average trend variation vector of

*M Nd(i) Ne(i) Nu(i)*

*Sd(i)* = *Sd(i*−1*)* + *(mid*\_*valuem* − *mid*\_*valuei*1*)* (13)

the *jth* subfactor *V*

= *( Sd(k) , Se(k) , Su(k) )*.

*S( j)*

*Nd(k) Ne(k) Nu(k)*

*Nd(i)* = *Nd(i*−1*)* + 1 (14)

*Se(i)* = 0 (15)

*Ne(i)* = *Ne(i)* + 1 (16)

Step 5: PSO-based optimum solution calculation: The optimal weight vector should be trained to calculate the final predicted re- sult. [Fig. 11](#_bookmark22) shows the detailed process.

First, the training dataset should be expressed as *T* = *(M, N)* ac-

cording to the following equation:

, ,

*R*1 *(t* − 1*) . . . Rm*+1*(t* − 1*) . . . Rn*+1*(t* − 1*)*

*. . .*

Otherwise, if *i*1 *< m*, then *Sa*(*i*) and *Na*(*i*) should be updated ac-

cording to the following equations:

*M* = ⎢, *R*1 *(t* − *i) . . . R*

*m*+1

*(t* − *i) . . . R*

*. . .*

, ,

*n*+1

*(t* − *i)* ⎥,*,*

*Sa(i)* = *Sa(i*−1*)* + *(mid*\_*valuem* − *mid*\_*valuei*1*)* (17)

*Na(i)* = *Na(i*−1*)* + 1 (18)

Assume that the SFTM of the current state belongs to group *i*. Then, the fuzzy-trend in the next state can be predicted by the following equation:

*R*1 *(t* − *p) . . . Rm*+1*(t* − *p) . . . Rn*+1*(t* − *p) R(t* − 1*)*

*. . .*

⎢, ⎥,

*N* = *R(t* − *i)* (24)

*. . .*

*R(t* − *p)*

*Sd(i)*

*Nd(i)*

*Su(i)*

*Nu(i)*

where ***M*** is a matrix formulated by the predicted values of *p*

*∆trend* =

*N*

*d(i)*

×

*Nd(i)*

+ *Ne(i)*

+ *Nu(i)*

+

*Nu(i)*

×

*Nd(i)*

+ *Ne(i)*

+ *Nu(i)*

# (19)

observations. ***N*** is the vector constructed by the corresponding ground truth values. *n* + 1 denotes the total number of granular spaces. *Rm*+1 *(t* − *i)* and *R(t* − *i)* denote the predicted values and

Step 4: Fuzzy-trend calculation for multigranular spaces: Divide the original dataset into *(n* + 1*)* granular spaces and predict the fuzzy-trend for each granular space based on the processes illus-

trated in [Fig. 10](#_bookmark21). In this study, the first granular space is defined as the principal factor, the second granular space is composed of the principal factor and one subfactor, and the (*n* + 1)th granular

space is composed of the principal factor and *n* subfactors. Here, *n*

denotes the number of subfactors.

In this procedure, the fuzzy-trend *∆m*+1 and corresponding pre- dicted value *Rm*+1 *(t)* of the (*m* + 1)th granular space can be calcu- lated using the following equations:

# 1

corresponding ground truth values for the *i*th observation in the

(*m* + 1)th granular space.

Next, the optimum weight vector should be determined by the

PSO method. In this procedure, *q* particles are generated within the *(n* + 1*)* granular spaces. The position of the *i*th particle is ex-

pressed as *Wi* = {*wi,*1*, wi,*2*, wi,*3 *, . . . . . . , wi,n*+1}, which should meet

the condition *wi, j* ∈ [0, 1] while *wi,*1 + *wi,*2 + *wi,*3 + *. . .* + *wi,n*+1 =

1. Similarly, the velocity of the *ith* particle is expressed as *Vi* =

{*vi,*1*, vi,*2*, vi,*3*, . . . . . . , vi,n*+1}, which should meet the condition *vi, j* ∈

[−1*,* 1]*, i* ≤ *j* ≤ *n* + 1. Both the position and velocity of each par-

ticle are initialized with random values. In this study, the fitness

function is defined as

*FWM* = 1 + Σ

*j*=1*,*2*,* *,m,r j >*0 *r j*

*r j*

# (20)

1

*f (i)* = *n* + 1 ×

*(N* − *M* × *Wi )* (25)

*i*

Σ

*FWs( j)* = 1 + Σ

*∆m*+1 = *FWM* × *∆M* + Σ

*j*=1*,*2*,* *,m,r j >*0 *r j*

*m* ¡ ¢

*FWs( j)* × *∆s( j)*

# (21)

(22)

where *f*(*i*) denotes the fitness value with 1 ≤ *i* ≤ *q*. Here, *q* repre- sents the number of particles generated in *(n* + 1*)* granular spaces. ***M*** and ***N*** are defined in the same manner as that in [Eq. (24)](#_bookmark23). ***W***

denotes the position vector.

*j*=1

*i*

*Rm*+1*(t )* = *R(t* − 1*)* + *∆m*+1 (23)

where *R(t* − 1*)* represents the raw data of the principal factor at the time point *t* − 1, *∆m* is the predicted fuzzy-trend of the prin- cipal factor, and *∆S*(*j*) is the predicted fuzzy-trend of the *j*th sub-

factor. In addition, *rj* is the correlation coeﬃcient between two fac- tors: (1) the average trend variation vector of the principal factor

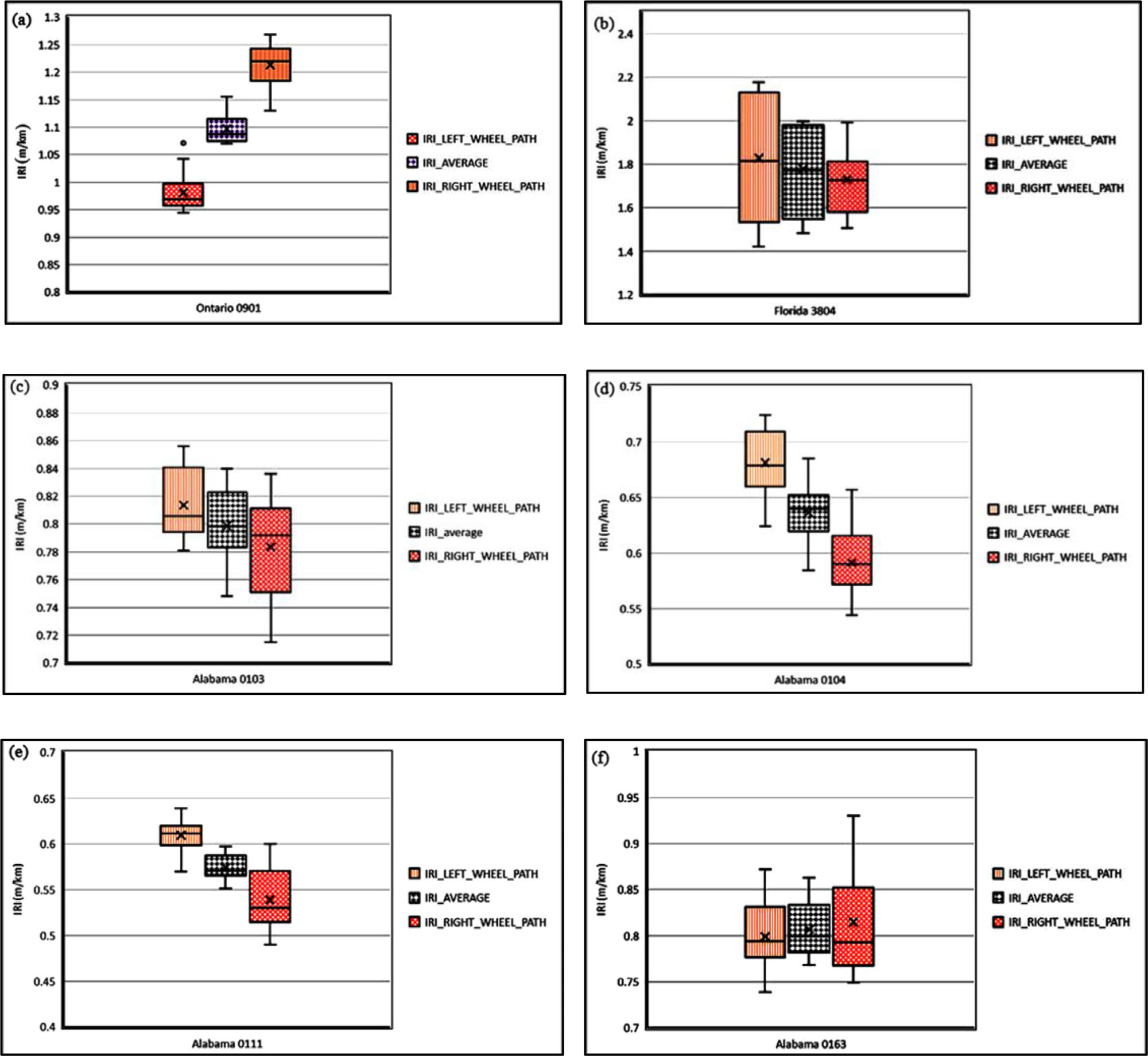
For an individual particle, after the update of each position and velocity, the corresponding fitness value should be calculated by [Eq. (25)](#_bookmark26). When the best fitness value is obtained, the forecasting trend and forecasted value can be calculated by [Eqs. (26)](#_bookmark28) and [(27)](#_bookmark30), respectively.

*n*

Σ

*∆z* = *wi*+1*∆i*+1 (26)

*i*=0



**Fig. 12.** Illustration of IRI distribution.

FR*(*t*)* = R*(t* − 1*)* + *∆z* (27)

where *∆i*+1 denotes the fuzzy-trend of the (*i* + 1)th granular space.

*∆z* denotes the corresponding forecasted fuzzy-trend. If *∆z >* 0,

the next state is considered to be in the ascending trend. Other- wise, it should be considered to be in the descending trend.

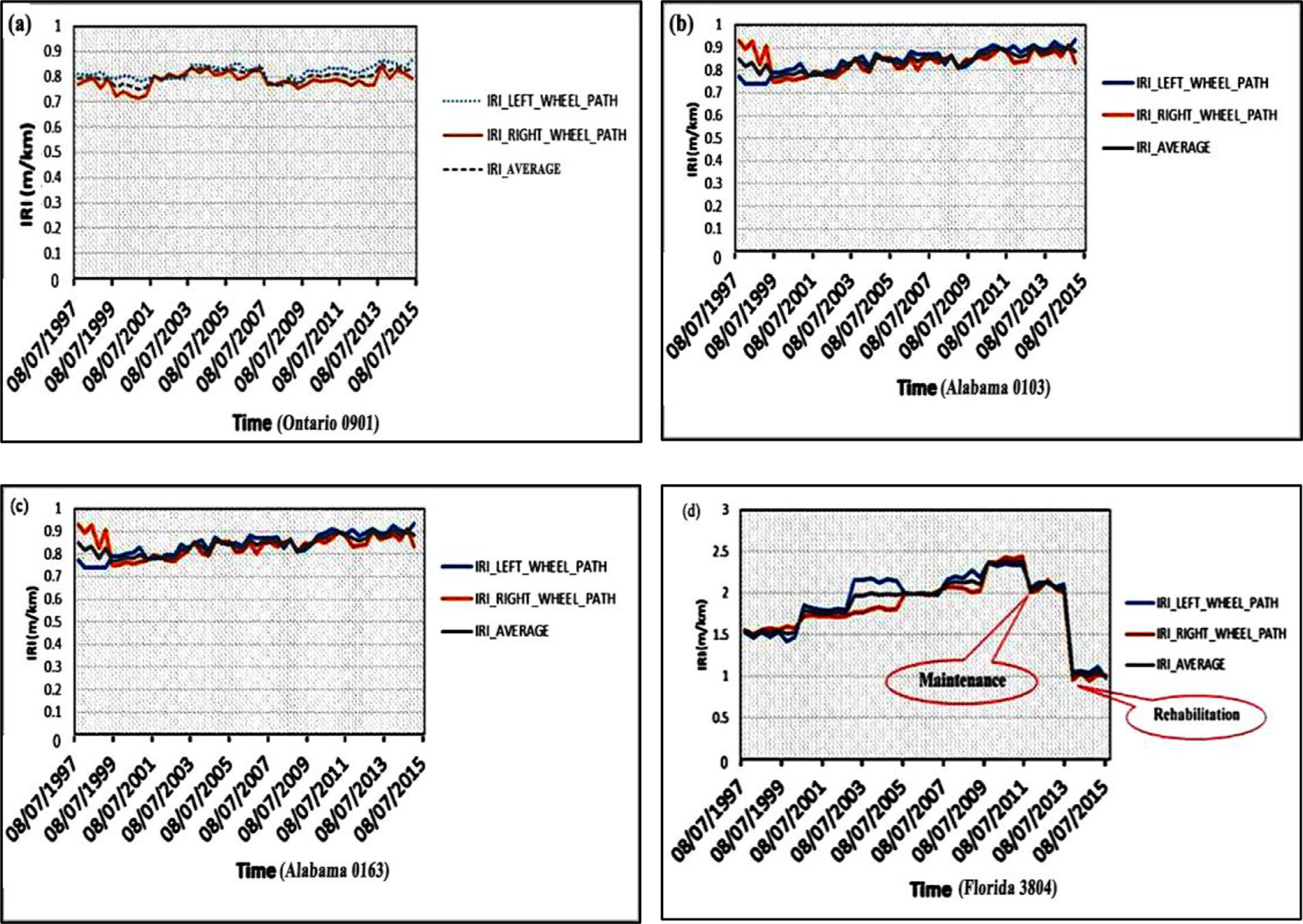
### Implementation details

* 1. *IRI data extraction and analysis*

IRI prediction experiments are performed using the proposed multigranularity fuzzy time-series method. Original IRI data are extracted from the LTPP database. Different from the traditional methods, which consider only a single factor when predicting IRI changes, the proposed method utilizes multiple factors. Specifi- cally, the average IRI values are selected as the principal factor, and the IRIs measured in the left and right wheel paths are used as the subfactors. The model predicts the average IRI in the next state with comprehensive consideration of the IRI values in the current state. Therefore, the results are more reliable.

Two significant phenomena are observed in the raw dataset:

1. The IRI distributions in different regions at the same time are significantly different from each other. This phenomenon can be seen in [Fig. 12](#_bookmark29). By comparing parts (a), (b), and (c) of this figure, it can be concluded that the global geographic locations (regions) of the pavements have significant influ- ence on the IRI values. [Fig. 12](#_bookmark29)(c)–(f) indicate that IRI vari- ance still exists between different road sections within the same global geographic location (region). However, the IRI difference between regions is more significant than that be- tween the different sections of the same region.
2. If no maintenance work is being carried out, the IRI tends to increase with the service time. This phenomenon can be seen in [Fig. 13](#_bookmark31), where parts (a), (b), and (c) represent the road sections that have not undergone maintenance opera- tions, according to the LTPP database. It is found that the overall trends of IRI changes agree with the expected in- creasing phenomenon, and only minor fluctuations are ob- served in section Alabama 0103 ([Fig. 13](#_bookmark31)(b)), which can be at- tributed to the errors introduced during data collection. Two



**Fig. 13.** IRI changes with time.

**Table 3**

Results of interval division.

|  |  |
| --- | --- |
| Factor | Intervals |
| Principal factor | [0.991, 1.216], [1.216, 2.182], [2.182, 2.268], [2.268, 2.653], [2.653, 2.682], [2.682, 2.917], [2.917, 3.075] |
| Subfactor 1 | [0.895, 1.137], [1.137, 1.791], [1.791, 1.825], [1.825, 1.926], [1.926, 2.006], [2.006, 2.07], [2.07, 2.076], [2.076, 2.146], [2.146, 2.205] |
| Subfactor 2 | [1.025, 1.304], [1.304, 2.561], [2.561, 2.615], [2.615, 3.207], [3.207, 3.366], [3.366, 3.732], [3.732, 3.845], [3.845, 4.027] |

repairing operations have been conducted in section Florida 3804 according to the records in the database. The first op- eration was a minor maintenance task, which reduced the IRI from around 2.5 to 2.0. After two years, a rehabilita- tion was carried out, which reduced the IRI to close to 1.0.

Then, intervals are generated based on the proposed ACT. [Table 3](#_bookmark32) presents the generated intervals.

(2) Fuzzy set formulation and FTRC establishment First, the fuzzy sets are formulated as follows:

This indicates that appropriate repairing work can improve

the overall performance of the pavement. Therefore, the de-

1

*A*1 =

*u*

1

0*.*5

+

*u*2

0

+ *. . .* + 

*u*7

velopment of effective IRI prediction models to forecast the

trends in potential IRI changes is very meaningful, and these models can be used to schedule cost-effective maintenance operations.

*A*1 =

0*.*5

*u*1

# 0

1

+

*u*2

*. . .*

# 0

0

+ *. . .* + 

*u*7

# 0

*A*1 =

*u*

1

+  + *. . .* +

*u*2 *u*7

# (28)

* 1. *IRI forecast experiments*

Because the IRI distribution is different for different regions, this research extracts datasets from 38 provinces/states in North America ([Fig. 14](#_bookmark33)). Thus, more than 20,000 data items are obtained to train the model, and hence, the robustness of the method can be assured. The experiments are discussed in detail below.

* + 1. Interval division using ACT

First, the original data are sorted in the ascending order. The average and standard deviations of each factor are calculated.

Next, FTRCs are built based on the SFTM. [Table 4](#_bookmark34) presents a part of the FTRCs for different factors. According to the fuzzy-trend re- lationships defined in [Table 1](#_bookmark17), Group 2 indicates equal trending be- tween two adjacent intervals. Therefore, the entire data presented in [Table 4](#_bookmark34) have equal trending, which indicates that the roughness conditions of these road sections have no significant variance be- tween each other. It should be noted that no groups are assigned to the values in the first two rows of [Table 4](#_bookmark34), which are indicated by a slash (/). These two values are not assigned to any groups because the group of current values is decided based on the eval- uation of the two values ahead of the current values.



**Fig. 14.** Regions considered for dataset selection.

**Table 4**

Part of FTRC result.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Average IRI | Fuzzy set | Group | Left-wheel-path IRI | Fuzzy set | Group | Right-wheel-path IRI | Fuzzy set | Group |
| 1.070999 | A1 | / | 0.964999 | B1,1 | / | 1.177000 | B2,1 | / |
| 1.080000 | A1 | / | 0.953999 | B1,1 | / | 1.20599 | B2,1 | / |
| 1.100999 | A1 | 2 | 0.968999 | B1,1 | 2 | 1.233000 | B2,1 | 2 |
| 1.105999 | A1 | 2 | 0.968999 | B1,1 | 2 | 1.243000 | B2,1 | 2 |
| 1.082999 | A1 | 2 | 0.975000 | B1,1 | 2 | 1.190999 | B2,1 | 2 |
| 1.080000 | A1 | 2 | 0.944000 | B1,1 | 2 | 1.215999 | B2,1 | 2 |
| 1.085999 | A1 | 2 | 0.952000 | B1,1 | 2 | 1.218999 | B2,1 | 2 |
| 1.103000 | A1 | 2 | 0.958000 | B1,1 | 2 | 1.248999 | B2,1 | 2 |
| 1.072000 | A1 | 2 | 0.962000 | B1,1 | 2 | 1.182000 | B2,1 | 2 |
| 1.088000 | A1 | 2 | 0.954999 | B1,1 | 2 | 1.220999 | B2,1 | 2 |
| 1.075999 | A1 | 2 | 0.992999 | B1,1 | 2 | 1.159000 | B2,1 | 2 |
| 1.070000 | A1 | 2 | 0.949000 | B1,1 | 2 | 1.190999 | B2,1 | 2 |

**Table 5**

Fuzzy-trend matrix for three factors.

* + - 1. Principal factor

Group Descending Equal Ascending

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Sum | Count |  | Sum | Count |  | Sum | Count |  |
| 1 | 0 | 0 |  | 0 | 1 |  | 1.375 | 1 |  |
| 2 | −1.717 | 2 |  | 0 | 30 |  | 1.520 | 3 |  |
| 3 | 0 | 0 |  | 0 | 4 |  | 0.733 | 2 |  |

* + - 1. Subfactor 1

Group Descending Equal Ascending

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Sum | Count |  | Sum | Count |  | Sum | Count |  |
| 1 | −0.107 | 1 |  | 0 | 1 |  | 0.996 | 5 |  |
| 2 | −0.725 | 3 |  | 0 | 20 |  | 1.483 | 4 |  |
| 3 | −0.623 | 3 |  | 0 | 5 |  | 0.068 | 1 |  |

* + - 1. Subfactor 2

Group Descending Equal Ascending

by the number 0. However, the IRI values for some road sections show ascending trends. These phenomena indicate that the rough- ness changing mechanisms in different road sections are signifi- cantly different even though they are distributed within limited ranges.

Thereafter, the fuzzy-trends of these three factors are calculated using [Eq. (19)](#_bookmark24). Here, the calculated fuzzy-trends are *∆principal* = 0*.*688*, ∆subfactor*1 = 0*.*127*, ∆subfactor*2 = 0*.*030. These results indi-

cate that all the fuzzy-trends tend to increase with time because

all of them are above 0.

(4) Fuzzy-trend forecast

The fuzzy-trends are constructed in multiple granular spaces. As discussed before, the first granular space includes only the average IRI values, the second granular space includes the average IRIs and one of the wheel-path IRIs, and the third granular space includes all the IRIs. Then, the fuzzy-trends and values of these granular

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Sum | Count |  | Sum | Count |  | Sum | Count |  | spaces can be calculated using [Eqs. (20)](#_bookmark25)–[(23)](#_bookmark27). |
| 1 | 0 | 0 |  | 0 | 0 |  | 1.029 | 2 |  | *∆*1 = 0*.*688*, ∆*2 = 0*.*4145*, ∆*3 = 0*.*248 |
| 2 | −0.148 | 1 |  | 0 | 31 |  | 1.211 | 3 |  |  |
| 3 | −0.323 | 1 |  | 0 | 4 |  | 0.656 | 1 |  |  |
|  |  |  |  |  |  |  |  |  |  | *R*1*(t)* = 3*.*684*, R*2*(t)* = 3*.*411*, R*3*(t)* = 3*.*244 |

1. Fuzzy-trend prediction for each factor

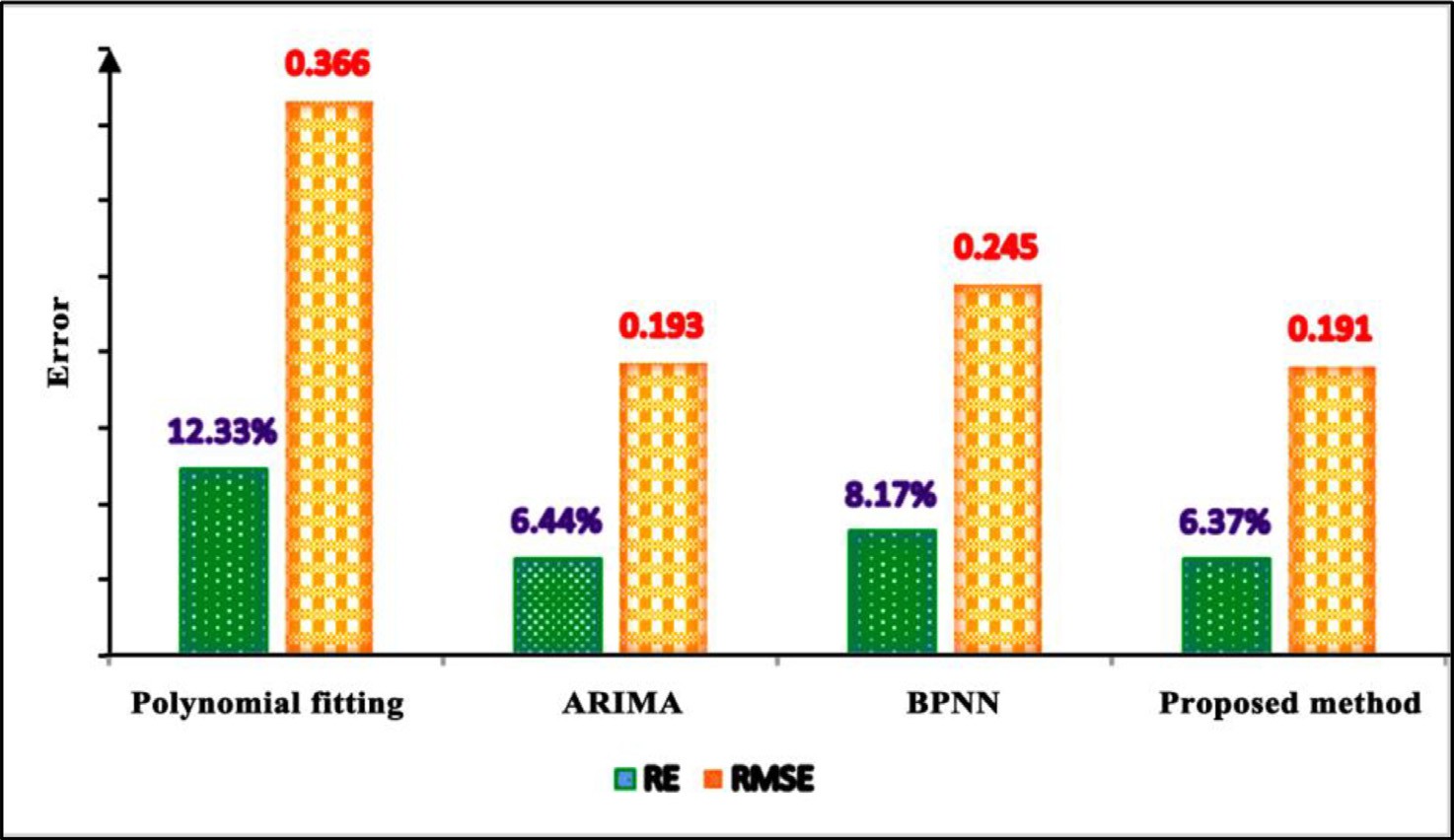
Three fuzzy-trend matrixes are generated according to the prin- ciple used to obtain the matrix in [Table 2](#_bookmark18). Accordingly, the ma- trixes presented in [Table 5](#_bookmark35) (a)–(c) are obtained. From these results, it can be inferred that few IRI values show descending trends, i.e., negative values. Most values remain constant, which is indicated

(5) Calculation of optimal weight vector based on PSO

In this study, 25 particles are used for model optimization. The optimum results are obtained after 200 iterations of parameter up- dating. Then, according to [Eqs. (8)](#_bookmark11) and [(9)](#_bookmark12), the optimal weight-

ing vector is calculated as *w* = *(*0*.*26*,* 0*.*11*,* 0*.*63*)*. Finally, the pre-

dicted IRI value is obtained based on [Eqs. (26)](#_bookmark28) and [(27)](#_bookmark30). The calcu-



**Fig. 15.** Performance of different methods.

lated values obtained from the experiment are

# *∆z* = 0*.*26 × 0*.*6875+0*.*11 × 0*.*4145+0*.*63 × 0*.*2477 = 0*.*3804

*F R(*t*)* = 2*.*996 + 0*.*3804 = 3*.*3764

Here, the calculated *∆z* is higher than 0, which indicates that the IRI tends to increase in the next state. *FR*, which is the fore- casted IRI value in the next state, is 3.3764.

### Results

To further verify the effectiveness of the proposed method, comparative experiments are performed. The polynomial fitting, autoregressive integrated moving average (ARIMA), and backprop- agation neural network (BPNN) methods are selected for IRI pre- diction, and the errors of these methods are compared with the results of the proposed method.

Two error indexes—the root mean squared error (RMSE) and relative error (RE)—are used to evaluate the feasibility of each method. These two indexes are calculated as follows:

, Σ*N* 2

for the pavement section Ontario 0901, Canada, to illustrate the entire prediction process. In addition, the IRI values for the pave- ment sections Florida 3804, Alabama 0103, and Alabama 0163 are extracted to perform comparative experiments.

Because times-series prediction relies highly on historical data, the data from 1997 to 2014 are collected for model training. Then, the trained model is used to predict the IRI value for the year 2015. The IRI values extracted from the LTPP database are considered as ground truth values to calculate the errors of the model. Thus, the IRI values used for model training are considered as the current state, whereas the IRIs for the year 2015 (the last year) are re- garded as the next state to be predicted.

Moreover, the prediction performance of the proposed method is compared with those of other widely used methods to verify the effectiveness and reliability of the proposed method. In summary, the following conclusions can be drawn from this research:

1. The original IRI data are divided into different length of inter- vals using the ACT. Then, three 3 × 6 fuzzy-trend matrixes are constructed to forecast the fuzzy-trend for each IRI factor. The

# RMSE =

*i*=1 *( pi* − *gti )*

*N*

# (29)

forecasting is achieved by defining the fuzzy sets and establish- ing FTRC.

1. To handle the uncertainty and achieve higher predicting accu-

*RE* = |*pi* − *gti* | × 100%*, i* = 1*,* 2*, . . . , N* (30)

## *gti*

where *N* denotes the number of the dataset. *pt* and *gtt* denote the

*ith* predicted IRI value and ground truth IRI value, respectively.

As indicated by the error distributions shown in [Fig. 15](#_bookmark36), the proposed method achieved the smallest error indexes (*RE* = 6*.*37%*, RMSE* = 0*.*191), followed by the ARIMA method (*RE* =

6*.*44%*, RMSE* = 0*.*193). The polynomial fitting method has the

least performance, which might be caused by the fact that tradi-

tional fitting methods are less capable of extracting comprehensive information from big datasets.

### Discussion

In this research, IRI data for 19 years (1997–2015) are extracted from the LTPP database for modeling. The IRI values were collected by the Federal Highway Administration. We select the IRI values

racy, IRI factors are divided into three granular spaces. On this basis, the IRI predictions are performed within each granular space. The optimal forecasted IRI values are calculated through PSO.

1. The proposed model employs the multigranularity method to provide highly reliable IRI predictions. By comparing the IRI prediction errors of the proposed method with those of the polynomial fitting, ARIMA, and BPNN methods, it is found that the proposed method has the smallest *RMSE* (0.191) and *RE* (6.37%) values, which indicate the effectiveness of the pro- posed approach.
2. The experimental results demonstrate that the proposed model can not only achieve the highest accuracy in IRI prediction but also provide detailed interpretative explanations of the trends in IRI changes (ascending, equal, or descending). These advan- tages are expected to provide significant benefits for protective pavement maintenance management.

### Data availability statement

The raw dataset used in this research can be downloaded from the LTPP InfoPave website, which can be accessed through the fol- lowing link: <https://infopave.fhwa.dot.gov/>.

### Funding

[This research is funded by Natural Science Foundation of Shaanxi Province [Grant number:](https://doi.org/10.13039/501100007128) [20](#_bookmark41)[17zdjc-23], Fundamental Re-](https://doi.org/10.13039/501100007128) search Funds for the Central Universities, Chang’An University [Grant number: [300102249301](#_bookmark42), [300102249306](#_bookmark42)], China.

### CRediT authorship contribution statement

**Wei Li:** Conceptualization, Methodology, Validation, Formal analysis, Writing - original draft, Writing - review & editing, Fund- ing acquisition. **Ju Huyan:** Conceptualization, Methodology, Valida- tion, Formal analysis, Data curation, Writing - original draft, Writ- ing - review & editing. **Liyang Xiao:** Methodology, Validation, For- mal analysis, Data curation, Writing - original draft, Writing - re- view & editing. **Susan Tighe:** Validation, Formal analysis, Writing - review & editing, Funding acquisition. **Lili Pei:** Validation, Formal analysis, Data curation, Writing - review & editing.

### Acknowledgements

Meanwhile, the authors would like to express sincere apprecia- tions to the LTPP staff for maintaining the Website and making the data so readily available. Moreover, the authors would like to thank the anonymous reviewers for providing their invaluable comments.

### Conflicts of interest

The authors declare no conflict of interest.

### References

Abdelaziz, N., Abd El-Hakim, R. T., El-Badawy, S. M., & Afify, H. A. (2018). Interna- tional roughness index prediction model for flexible pavements. *International Journal of Pavement Engineering*, 1–12. doi:[10.1080/10298436.2018.1441414](https://doi.org/10.1080/10298436.2018.1441414).

[Albuquerque, F., & Núñez, W. (2011). Development of roughness prediction mod- els for low-volume road networks in Northeast Brazil. *Transportation Research* *Record: Journal of the Transportation Research Board, 2205*, 198–205.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0002)

Cebon, D. (1993). *Interaction between heavy vehicles and roads* SAE Technical Paper. doi:[10.4271/930001](https://doi.org/10.4271/930001).

[Chen,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0004) [S.-M. (1996). Forecasting Enrollments Based on Fuzzy Time Series. *Fuzzy Sets and Systems, 81*(3), 311–319.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0004)

[Chen,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0005) [S.-M., & Jian, W.-S. (2017). Fuzzy forecasting based on two-factors sec- ond-order fuzzy-trend logical relationship groups, similarity measures and Pso techniques. *Information Sciences, 391*, 65–79.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0005)

[Chen,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0006) [S.-M., Manalu, G. M. T., Pan, J.-S., & Liu, H.-C. (2013). Fuzzy forecasting based on two-factors second-order fuzzy-trend logical relationship groups and par- ticle swarm optimization techniques. *IEEE Transactions on Cybernetics, 43*(3), 1102–1117.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0006)

[Chen,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0007) [S.-M., & Tanuwijaya, K. (2011). Fuzzy forecasting based on high-order fuzzy logical relationships and automatic clustering techniques. *Expert Systems with Applications, 38*(12), 15425–15437.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0007)

[Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. (2003). Arima models to predict next-day electricity prices. *IEEE Transactions on Power Systems, 18*(3), 1014–1020.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0008)

[George, K. (2000). *Mdot pavement management system: prediction models and feed- back system*. Mississippi: Dept. of Transportation.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0009)

[Georgiou, P., Plati, C., & Loizos, A. (2018). Soft computing models to predict pave- ment roughness: A comparative study. *Advances in Civil Engineering, 2018*, 1–8.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0010)

[Hao,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0011) [C., Wang, Y., & Tuo, J. (2015). A novel particle swarm optimization algorithm with intelligent weighting mechanism. In *Information science and control engi- neering (ICISCE), 2015 2nd international conference on* (pp. 45–49). IEEE.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0011)

[Huarng,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0012) [K. (2001). Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy Sets and Systems, 123*(3), 387–394.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0012)

[Jackson, N., Puccinelli, J., & Engineers, N. C. (2006). Long-term pavement perfor- mance program (Ltpp) data analysis support: National pooled fund study Tpf-5](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0013)

[(013). *Effects of multiple freeze cycles and deep frost penetration on pavement per- formance and cost*. Turner-Fairbank Highway Research Center.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0013)

[Kargah-Ostadi, N., Stoffels, S., & Tabatabaee, N. (2010). Network-level pavement roughness prediction model for rehabilitation recommendations. *Transportation Research Record: Journal of the Transportation Research Board, 2155*, 124–133.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0014)

[Kuo, I.-H., Horng, S.-J., Kao, T.-W., Lin, T.-L., Lee, C.-L., & Pan, Y. (2009). An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. *Expert Systems with Applications, 36*(3), 6108–6117.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0015)

[Lu,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0016) [P., & Tolliver, D. (2012). Pavement pre-and post-treatment performance models using Ltpp data. *Journal of the Transportation Research Forum, 51*(3), 67–81 Cite- seer.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0016)

[Oﬃcials,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0017) [T. (2008). *Mechanistic-empirical pavement design guide: A manual of practice*.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0017)

[AASHTO.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0017)

[Perera, R., & Kohn, S. D. (2001). *Ltpp data analysis: factors affecting pavement smooth- ness*. Transportation Research Board, National Research Council.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0018)

[Peshkin, D. G., Hoerner, T. E., & Zimmerman, K. A. (2004). *Optimal timing of pavement preventive maintenance treatment applications*. Transportation Research Board.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0019)

[Qian,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0020) [J., Jin, C., Zhang, J., Ling, J., & Sun, C. (2018). International roughness index pre- diction model for thin hot mix asphalt overlay treatment of flexible pavements. *Transportation Research Record* 0361198118768522.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0020)

[Rahim, A. M., Fiegel, G., Ghuzlan, K., & Khumann, D. (2009). Evaluation of interna- tional roughness index for asphalt overlays placed over cracked and seated con- crete pavements. *International Journal of Pavement Engineering, 10*(3), 201–207.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0021)

[Reuter, U.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0022) [(2007). *Uncertainty forecasting in engineering*. Springer](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0022). [Sets, F.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0023), [&](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0023) [Zadeh, L.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0023) [(1965). Inform. *Control, 8*, 338–353](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0023).

[Shirazi, H., Ayres, M., Speir, R., Song, W., & Hall, G. (2010). Confidence level and ef- ficient sampling size of roughness measurements for pavement management in Maryland. *Transportation Research Record: Journal of the Transportation Research Board, 2153*, 114–120.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0024)

[Song,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0025) [Q., & Chissom, B. S. (1993). Fuzzy time series and its models. Fuzzy Sets and Systems*, 54*(3), 269–277.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0025)

[Wang, G.-Y., Zhang, Q.-H., Ma, X.-A., & Yang, Q.-S. (2011). Granular computing mod- els for knowledge uncertainty. *Journal of Software, 22*(4), 676–694.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0026)

[Yao, Y. (2000). Granular computing: basic issues and possible solutions. In *Proceed- ings of the 5th joint conference on information sciences* (pp. 186–189).](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0027)

[Yao, Y.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0028) [(2005). Perspectives of granular computing. *GrC, 5*, 85–90](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0028).

[Yu,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0029) [H.-K. (2005). A refined fuzzy time-series model for forecasting. *Physica A: Statis- tical Mechanics and its Applications, 346*(3-4), 657–681.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0029)

[Zadeh,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0030) [L. A. (1996). Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy*](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0030)

[*Systems, 4*(2), 103–111.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0030)

[Zhao,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0031) [Z., & Guo, Z. (2013). Prediction of asphalt pavement international roughness index (Iri) by combined approach of empirical regression and Markov. In *Icte 2013: Safety, speediness, intelligence, low-carbon, innovation* (pp. 2106–2113).](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0031)

Zhou, G., Wang, L., & Lu, Y. (2008)., International roughness index model enhance- ment for flexible pavement design using Ltpp data.

[Zimmerman,](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0032) [K., & Peshkin, D. (2003). Pavement management perspective on in- tegrating preventive maintenance into a pavement management system. *Trans- portation Research Record: Journal of the Transportation Research Board, 1827*, 3–9.](http://refhub.elsevier.com/S2590-1885(19)30006-X/sbref0032)