Early Warning and Monitoring System of the Economic Situation in Real Estate Market

Dongmei Han
School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, P.R.China
Shanghai Financial Information Technology Key Research Laboratory, Shanghai 200433, P.R.China
Email: handongmei19610320@gmail.com

Yonghui Dai and Zhiyi Zhang
School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China
Email: dyh822@163.com, zzy2020@gmail.com

Abstract—Currently, many scholars have theoretically studied the early warning index of real estate market, but few have focused on the implementation of the system. Our study provides a system for early warning and monitoring the economic situation in real estate market. We construct a boom index for China's real estate market based on the state space model. In particular, we provide users with intuitive graphical interface, which helps regulators distinguish the economic situation of real estate market in a scientific and convenient way. In addition, the system implementation includes some key technologies, such as model-solving, conversion and fusion among multi-data sources, cross-language invoking and hybrid programming.

Index Terms—Early warning and monitoring system, state space model, boom index, data source integration

I. INTRODUCTION

The early warning and monitoring of real estate situation have attracted scholars’ great attention. Peng(2002) proposed that the establishment of early warning system should be constructed the following two aspects: the transverse frame and the longitudinal frame. The transverse frame includes three subsystems: the sampling survey system, the economic analysis system and the early warning signal system. The longitudinal frame consists of the monitoring and warning system based on country, region (province) and city levels. These frames can make the early warnings system solid [1]. Yang, et al. (2006) provided a method to screen indexes based on clustering analysis and non-parametric tests, and applied it in the empirical analysis of the early warning of Shanghai real estate market diffusion index [2]. In addition, the artificial neural network method and econometric models have been widely applied to forecast and analyze the economic situation in real estate market [3]-[4]. Up to now, many scholars have theoretically studied the indexes of early warning system [5]-[12], but few have focused on the implementation of the system [13]-[14], in that there are some difficulties to develop early warning system.

1. Heterogeneous data lead to difficulty in data fusion. Data of government statistics departments, financial institutions, enterprises and institutions is often managed dispersely, and is mostly distributed in heterogeneous data platform, which possesses problems of data format inconsistent and leads to difficulties of data fusion.

2. To be universal, this system should be a flexible business support system [15]. Due to the requirement of many econometric model can be flexibly invoked, hybrid programming is essential, which involves complicate technology.

3. How to show users the economic development of the real estate market situation in an intuitive way is also a big problem.

Taking full considerations of the above problems, the system in the paper gives a good solution.

II. SYSTEM ARCHITECTURE

Taking into account the universal, this system architecture is designed with three layers: resource service layer, application layer and user layer. The econometric model of the resource service layer is encapsulated into dll(dynamic link library), then be invoked in hybrid programming. Application layer includes five functional subsystems: data processing subsystem, index management subsystem, boom index subsystem, the monitoring and warning subsystem and...
auxiliary management subsystem. Main framework of system as shown in Fig. 1.

III. SYSTEM DESIGN

A. Subsystem Design

This system include five subsystem, the functional module of each subsystem as shown in Table I.

<table>
<thead>
<tr>
<th>Number</th>
<th>Function module</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data Processing Subsystem</td>
<td>Construct a year-on-year growth indicators</td>
</tr>
<tr>
<td>2</td>
<td>Indicators System Management Subsystem</td>
<td>Indicator System Management</td>
</tr>
<tr>
<td>3</td>
<td>Boom Index Subsystem</td>
<td>Boom indicator screening methods</td>
</tr>
<tr>
<td>4</td>
<td>Monitoring and Warning Subsystem</td>
<td>Calculate the turning point</td>
</tr>
<tr>
<td>5</td>
<td>Auxiliary Management Subsystem</td>
<td>System parameter settings</td>
</tr>
</tbody>
</table>

Some features of the module are as follows:

1. Data processing

In data processing module, it can do various pre-processing based on the different index type, which means that it can distinguish annual data, quarterly data, monthly data, as well as the cumulative value, and the month value type.

2. Boom indicator screening methods

The index group would be divided into 3 parts according to the time sequence among them: prior index group, consistent index group and lagging indicator group. The prior indicator is used to predict the future trend, the consistent indicator reflect the current status, and the lagging indicator is used to Ex-post verification [16]. The system provides methods include the time-difference correlation analysis, K-L information method, cluster analysis [17] and principal component analysis etc [18].

3. Synthetic boom index

In the aspect of indicating the boom movements of the real estate market through the synthesis of various economic indicators boom index. This system not only provides traditional diffusion index and composite index [19], but also provides SWI boom index based on state space model.

4. 3σ monitoring and early warning method

For monthly data, we divides the cordon warning domain into the normal range, partial heat range, overheating range, colder range and too cold range and present the result in graphic demonstration.

B. Building Boom Index based on State Space Model

The state space model is a dynamic time-domain model, and takes implicit time as independent variable. It is often used to estimate unobservable time-variables: rational expectations, measurement error, long-term revenue and unobservable factors [20]-[22]. Dynamic system using state space model has two main advantages: Firstly, unobservable variables can be observed and estimated when merging into the observable models. Secondly, taking advantage of strong calculation function of Calman filter, the state space model has the very high precision, for it is realized by constantly recursive approximation to the given value [23]-[24].

This system uses a state space model to construct a real estate boom index. In fact, there are many variables are unobserved in the early warning and monitoring system. These unobservable variables are called state vector, which reflects the true economic state of the real estate market. The boom index structure based on the state space model opens up a new way to index synthesis in the real estate field.

1. Building state space model

The state space model consists of two equations: one is the measurement equation, another is the state equation. The measurement equation reflects the relation between the observation sequence and system status, and the state equation reflects the transition from the current state to next state.

We construct a synthetic boom index state space model using time series of five coincident indicators: value of commercial housing completed, commercial housing area...
completed this year, commercial housing sales amount, commercial housing sales area, self-raised funds of commodity house investment. Let $y_t$ be the value of 5 seasonally adjusted economic variables from which change trend is removed at time $t$, so, it is a $5 \times 1$ Vector and is observable. $c_t$ is the true value of the boom index at time, and it is a scalar, denoted by $c = [c_1, c_2, ..., c_T]$. $c_t$ is the SWI boom index, and it is unobservable. Based on decomposition principle of time series, $y_t$ is divided into two parts, one is recursive element, the other is irregular element. So, the time series can be expressed in the following formula:

$$y_t = x c_t + u_t$$  

Equation (2) is known as measurement equation, which doesn’t contain observation error. Then, regarding $c_t$ and $u_t$ as stochastic process, $c_t$ is AR(p) (p-step autoregressive process), each component of $u_t$ is AR(r)(r-step autoregressive process), so, the expressions are like this:

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \cdots + \phi_p c_{t-p} + \eta_t$$

$$u_t = \theta_{1t} u_{t-1} + \theta_{2t} u_{t-2} + \cdots + \theta_{r_t} u_{t-r} + \epsilon_t$$

Assume

$$\eta_t \sim N(0, \sigma^2)$$

$$\epsilon_t \sim N(0, \sigma^2 H)$$

And the two above is independent, here

$$H = \text{diag}(h_1, \cdots, h_k)$$

(3) and (4) are state equations. Parameters $\gamma, \phi, \theta, \sigma^2, H$ are hyper parameters to be estimated, and $(p, q, r)$ is jitter time of the model.

From the above model construction process, we can see the delay order $p, q, r$ determines the basic shape model, this paper reference literature, when $p=2$, order is only affected by two, reflecting the economic significance is not very significant, so $p = 3$ as the initial order; $q, r$ directly decided the measurement equation and the equation of transfer of a number of parameters to be estimated, so the order starts from 2. Therefore, the initial state space model of the delay of order $(3, 3, 2)$, the state space model consistent with 5 indexes are follows.

**Measurement Equation:**

$$
\begin{align*}
Y_{1t} &= \gamma_1 c_t + \gamma_{12} c_{t-1} + \gamma_{13} c_{t-2} + u_{1t} \\
Y_{2t} &= \gamma_{21} c_t + \gamma_{22} c_{t-1} + \gamma_{23} c_{t-2} + u_{2t} \\
Y_{3t} &= \gamma_{31} c_t + \gamma_{32} c_{t-1} + \gamma_{33} c_{t-2} + u_{3t} \\
Y_{4t} &= \gamma_{41} c_t + \gamma_{42} c_{t-1} + \gamma_{43} c_{t-2} + u_{4t} \\
Y_{5t} &= \gamma_{51} c_t + \gamma_{52} c_{t-1} + \gamma_{53} c_{t-2} + u_{5t}
\end{align*}
$$  

**(8)**

**State Equation:**

$$
\begin{align*}
c_t &= \phi_1 c_{t-1} + \phi_2 c_{t-2} + \phi_3 c_{t-3} + \eta_t \\
u_{1t} &= \theta_{11} u_{1t-1} + \theta_{12} u_{1t-2} + \epsilon_{1t} \\
u_{2t} &= \theta_{21} u_{2t-1} + \theta_{22} u_{2t-2} + \epsilon_{2t} \\
u_{3t} &= \theta_{31} u_{3t-1} + \theta_{32} u_{3t-2} + \epsilon_{3t} \\
u_{4t} &= \theta_{41} u_{4t-1} + \theta_{42} u_{4t-2} + \epsilon_{4t} \\
u_{5t} &= \theta_{51} u_{5t-1} + \theta_{52} u_{5t-2} + \epsilon_{5t}
\end{align*}
$$  

**(9)**

In (8), $y_{1t}, y_{2t}, y_{3t}, y_{4t}, y_{5t}$ represent five indicators: value of commercial housing completed, commercial housing area completed this year, commercial housing sales amount, commercial housing sales area, self-raised funds of commodity house investment.

From the above (8) – (9), we can see, $\gamma$ contains $5 \times 3 = 15$ parameters, $\phi$ contains 3 parameters, and $\theta$ contains $5 \times 2 = 10$ parameters. In addition, variance items contains 6 parameter. So, There are total of 34 parameters to be estimated consist of super parameter set.

2. SWI boom index based on state space model

Boom index is an effective method when government and enterprise research on economic cycle fluctuation, evaluate and analyze macro economic development state. James H. Stock(American) and Mark W. Watson(American) put forward a new boom index structure method in 1988, they think there is an unobservable variables represent the general state of the economy, and its fluctuation is the real boom cycle, the unobservable variables, known as type Stock-Watson boom index, referred to as SWI [20]. Based on (8) – (9), this paper established state space model with 5 coincident indicators chosen by SSPACE of Eviews7.0 software. Time series is growth rate series after seasonal adjustment. In order to avoid fluctuations in the individual indicators, the paper also normalized the indicators series. The expression of normalization is like this, $X' = (X - \bar{X}) / S_x$, here, $\bar{X}$ is average of raw data and $S_x$ is standard deviation of raw data. We use the SSPACE tool to estimate the set hyperparameters with logarithmic likelihood estimation, then the state vector
using Calman filtering solution. In the case of the parameters are not being given the initial value, model-solving obtain the state variables \( c = (c_1, \cdots, c_T) \), named SWI index as shown in Fig. 2.

As you can see, the SWI index generally reflects the fluctuation situation of China's real estate market in 1999-2011 years. From the 1998 housing reform, China's real estate has entered a stage of rapid development. Boom peaks appear four times, respectively, at the end of 1999, the end of 2002, the end of 2005 and the end of 2009. In 2003, the state carry out macro-control of the real estate market, and are subject to the impact of the SARS epidemic, making the real estate boom lower level. In 2004, the state has adopted polices of "tightening", "control of land supply"; the development of the real estate market is stable in the low. In 2005, because the relationship between supply and demand in the market is tight, commercial housing sales area and sales of commercial housing are year-on-year rapid growth, real estate boom level arrive at the highest level in history. The global financial crisis of 2008 caused the real estate market boom levels dropped to the lowest point in history, in 2009, affected by the policy to rescue the market, real estate turnover of warmer. Since 2010, China promulgated a series of macro-control policies and restrictions on property measures, so that the development of the real estate market more rational.

From the above analysis, the fluctuation tendency of using state space model SWI index synthetic fit the realistic situation, which means the synthesis method is effective.

IV. SYSTEM IMPLEMENTATION

A. B/S Architecture based on Multiple Data Sources

Based on multiple data sources, B/S architecture is composed of three layers as shown in Fig. 3. In the presentation layer, various clients connect to the Internet through the browser, realize human-computer interaction in the information system. The business logic layer, consisting of the web server with multiple data sources, to meet the needs of a variety of data format conversion and integration. Besides, The back-end database server uses relational database. In this architecture, the web server responses to the request from the clients, and implements the integration and transformation of multiple data sources, including the import and export of TXT format data, the import and export of Excel format data sources, XML data transformation, and graphical output.

B. Development Platform and Runtime Environment

Operating system: Windows 7 Home Basic
Development Tool: Microsoft Visual Studio 2008
Database: SQL Server 2005 Enterprise Edition
Programming language: ASP.NET (C#)
WEB Server: IIS 7.0
Model library software: Matlab R2010b, Eviews 7.0
Image output: Amcharts Flash
Middleware interface technology: COM Automation, Excel VBA

C. Cross-language Call the Model Library

1. Cross-language calls Matlab and Eviews model library architecture

Take advantage of statistical analysis software Matlab and econometric software Eviews in the boom index module and early warning module, we can overcome the shortcomings of calculation and model. The system constructs some models such as diffusion index, composite index, time-difference correlation analysis model in Matlab, and builds the X12 seasonal adjustment model, the state space model in Eviews software. Based on the package for Matlab and Eviews software, the web server calls and integrates the econometric model. The call architecture as shown in Fig. 4.
The model library system of the structure server is transparent to the users. Users only need a browser can call server-side Matlab and Eviews software models and features without the need to install any other software. Matlab model library package include diffusion index, composite index, difference correlation analysis and other function code, All code is packaged compiled into a dynamic link library, deployed on an web server and application server. The same as to eview model library package.

This architecture simplifies the client load, led the server has high scalability, and achieve resource sharing and integration.

2. Examples based on the package for Eviews software

Take the X12 seasonal adjustment as an example, the system demonstrates to call eviews software. First, add a reference to the Excel COM and the corresponding table space in the environment of C#, after the main program can create a new Excel Application object, then open an existing workbook file data1, and call the RunMacro method to run the X12 seasonal of VBA model in data1, then pass syntax string as a parameter to VBA according to the Eviews seasonal adjustment command line syntax, code is showed in Fig. 5.

![Figure 5. Original sequence before seasonal adjustment](image)

Then, the VBA module executive name is X12 seasonal macro, the macro call the Manager class of Eviews COM to get a Application app object for an instantiation, it will be able to receive the seas parameters by running the app.run command, and make a sequence after seasonal adjustment return to the Excel table, and then through the OLEDB back to the C# program, VBA code is showed in Fig. 6.

```csharp
using Microsoft.Office.Interop;
using Excel = Microsoft.Office.Interop.Excel;
using System.Runtime.InteropServices.ComTypes;
private void RunMacro(object oApp, object[] oRunArgs)
{
    oApp.GetType().InvokeMember("Run",
            null, oApp, oRunArgs);
}
```

Finally, the obtained data table by drawing module output. Original series and seasonal adjustment of the sequence as shown in Fig. 7-Fig. 8:

![Figure 7. Original sequence before seasonal adjustment](image)

![Figure 8. Adjusted the original sequence and cycle sequence](image)

By calling the packaged X12 seasonal adjustment module based on Eviews software, the system gives intuitive graphical interface, which helps user judge the economic situation of China’s estate market in a convenient way.

D. Graphical Display the Results

In the boom index module, select the state space model SWI index, set delay order number for (5, 3, 2), and choose to estimate parameters of the initial $\gamma_1$ value of
0.1-1. The system calls the packaged state space model, get ten SWI index synthetic results, As shown in Fig. 9.

From Fig. 9 shows, we can see SWI index is the smallest fluctuation when the $\gamma_1$ value is equal to 1.

V. KEY TECHNOLOGIES

A. Transformation and Integration of Multiple Data Sources

Data Table is an intermediate form and convenient to process by the procedure. In the script language, we adopt ADO.NET (Active X Data Object) to connect the server and back-end database, and exploiting SQL Connection object to realize write-in and write-out between Date Table and date of the relational database.

We access the Txt data source through the file stream input and output, quote the System.IO table space, declare the File Stream objects, and transfer the Data Table data and Txt format by using Stream Reader and Stream Writer method of the file stream object to realize import and export.

We use OLEDB to access the Excel form data to realize import and export between Data Table and Excel form data. In order to output the final image output by imaging technology, we should transfer the Data Table to XML standard transmission format data, declare the XML document object under the XML space, and create XML format files with the contents of Data Table data table through XML Element method. The multiple data sources transformation and integration design scheme as shown in Fig. 10.

B. Conversion and Transmission of Data with Matlab, Eviews

The core of this system is to achieve the package and call of Matlab and Eviews software. Because the processing math model and format of these two software is completely different from server-side Data Table, we need to design data transformation and delivery methods between Web server and these two software.

Data calculated by Matlab must be two-dimensional digital matrix. So the Data Table must be transformed to two-dimensional matrix before the data transmitted to and processed by Matlab. The method is to transform Data Table into intermediate format of two-dimensional Array by C# program, and then by the introduction of the MathWorks, MATLAB. NET, Arrays table space, transform the two-dimensional array into MWNumericArray digital matrix which can be identified by Matlab, transmit to the Matlab model for processing. After processing the results should be transformed into Data Table by the same way.

Eviews internal data are stored in workfile and processed. The transformation procedure is relatively complex, and is realized with Excel as the middleware. Because the C# language itself cannot access Eviews software, we must create a server component to access through the COM interface. Firstly the data is transformed to the Excel through the OLEDB, secondly we can call the VBA in the Excel workbook through the COM component, which also uses the COM Automation component provided by Evies software. And Finally, after Eviews processing, the data return to the Excel Sheet and back to the server through the OLEDB. The related data transformation and transition design as shown in Fig. 11.

C. The key Technologies of Mixed Calling the Back-end Database

1. Dynamically building table

In conventional database design, database table is fixed. The monitoring and early-warning system in this paper needs to confirm the number of index dynamically, and concern the differences of each index beginning and finishing times, and the annual, quarterly, monthly data frequencies. Thus the conventional data structure design cannot meet the demand.

© 2013 ACADEMY PUBLISHER
Therefore, we use the trigger to create dynamical table. When executing the SQL statement dynamic to create table, the table name corresponds to the primary key of index in the basic information table. After adding a new index, there is a piece of new basic information in the index information table, and a new data table dynamically created; by maintaining multiple tables synchronization via the trigger, we realize the back-end database maintained dynamically, and the zero vacancy rate of storage resources. It is the most optimized design.

2. Data Alignment

Data alignment, in other words, equivalent to select the public time interval observed values in multiple time series before the index synthesis. So the design of index data table must contain the date list, and the annual, quarterly and monthly data has respectively date format, such as "year 2000", "year 2000 1st quarter", and "year 2001 Jan". Thus when jointing the tables, all index observed values of public time interval of the same type indicator is drawn by inner join way.

According to the above database design idea, we make the data model and the entity relationship diagram (E-R diagram), as shown in Fig. 12.

![Diagram](image)

Figure 12. Data alignment relation

In Fig. 12, index table is created dynamically, and it includes three properties: ID, CurrentDate and Value. This index table named Index N, and N value taken from Table MetaIndex’s column IndexID.

VI. DISCUSSION AND CONCLUSION

The real estate early-warning and monitoring system of this paper is universal and convenient. Firstly, it can be flexibly extended by other econometric modules. Secondly, it can show the fluctuation of the real estate market in an intuitive graphical way. What’s more, it can offer early warning timely and accurately to meet the demands of the regulars and scholars, and provide more convenient and objective reference to related departments when making the macro decisions.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China (No.41174007) and High Level Academic Research Program (No. 2012DFRRCG0D02) of Financial Research Center, Fudan University.

REFERENCES


**Dongmei Han** 1961-, China. She is full Professor and Doctoral Advisor at the School of Information Management and Engineering, Shanghai University of Finance and Economics, China. She received her Ph.D. in 2002 from Quantitative Economics in Jilin University, China. Currently, Her research interests include finance engineering, management science and Engineering.

**Yonghui Dai** 1975-, China. He is current a Ph.D. student at the School of Information Management and Engineering, Shanghai University of Finance and Economics, China. He received his Master degree in Software Engineering in 2006 from Fudan University, China. His research interests include financial information technology, behavioral finance and mobile information technology integration.

**Zhiyi Zhang** 1987-, China. He received his Master degree in Management Science and Engineering in 2011 from Shanghai University of Finance and Economics, China.