



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 9, September 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.524

9940 572 462

6381 907 438

ijrset@gmail.com

www.ijrset.com

Study on Location Selection Method for Multinational Service Company based on Multiple Models

Min Duan

Teaching Department of Public Courses, Hunan Communication Polytechnic, Changsha, China

ABSTRACT: This article aims to help the multinational service company to investigate trends of global languages and location options for new offices, based on 2018 Mathematical Contest in Modeling (MCM) Problem. The logistic curve model was adapted to predict the trend of future population change and the example of China was used to introduce the modeling process in detail. The changing patterns of the population were associated with many uncertainties. The grey forecast model was proposed to predict what will happen to the numbers of native speakers and total language speakers in the future. The results showed that the list of top 10 most-used languages has changed many times in the past. English will replace Chinese as the new most-spoken language in the world over the next 50 years, and the number of Arabic and Spanish users was growing rapidly with a trend of surpassing Chinese. As for the six new international offices that the client company wants to establish, in the short term, six optimal locations are obtained based on the K-means clustering method of the latitudes and longitudes of global cities. The choice of long-term international office address needs to consider not only the geographical location, but also the gross domestic product (GDP) of the city or country, so as to promote the business and obtain better returns. Finally, the locations of new international offices were optimized according to GDP and geographical location.

KEYWORDS: Mathematical modeling; Grey prediction; K-means clustering; Simulated annealing algorithm.

I. INTRODUCTION

There are currently about 6,900 languages spoken on Earth. About half the world's population claim one of the following ten languages (in order of most speakers) as a native language: Mandarin (incl. Standard Chinese), Spanish, English, Hindi, Arabic, Bengali, Portuguese, Russian, Punjabi, and Japanese. However, much of the world's population also speaks a second language. When considering total numbers of speakers of a particular language (native speakers plus second or third, etc. language speakers), the languages and their order change from the native language list provided. The total number of speakers of a language may increase or decrease over time because of a variety of influences to include, but not limited to, the language(s) used and/or promoted by the government in a country, the language(s) used in schools, social pressures, migration and assimilation of cultural groups, and immigration and emigration with countries that speak other languages. Moreover, in our globalized, interconnected world there are additional factors that allow languages that are geographically distant to interact. These factors include international business relations, increased global tourism, the use of electronic communication and social media, and the use of technology to assist in quick and easy language translation.

A large multinational service company, with offices in New York City in the United States and Shanghai in China, is continuing to expand to become truly international. This company is investigating opening additional international offices and desires to have the employees of each office speak both in English and one or more additional languages. The Chief Operating Officer of the company has hired your team to investigate trends of global languages and location options for new offices.

In order to solve the 2018 Mathematical Contest in Modeling (MCM) Problem B [1], some intelligent algorithms were applied. The logistic curve model was adapted to predict the trend of future population change. The grey prediction model was proposed to predict what will happen to the numbers of native speakers and total language speakers in the next 50 years. The K-means clustering method was used to determine the location of new international offices, which considered space coverage as large as possible, and the gross domestic product (GDP) of the city or country was considered to optimize the location of offices.

II. POPULATION PREDICTION METHOD

Since the number of language users is closely related to population base, it is necessary to establish a suitable model to describe the population trend over time. The population prediction model was established based on the logistic curve, and China was taken as an example to fit the prediction model, in which the data comes from the website of the National Bureau of statistics of the people's Republic of China.

Table 1 showed the population of China in the past years, where the variable t represented the time serial number from 1951, X represented the population and the unit was ten thousand.

Table 1. The population of China in the past years

year	t	X/ Ten thousand	year	t	X/ Ten thousand
1951	1	56300	1984	34	104357
1952	2	57482	1985	35	105851
1953	3	58796	1986	36	107507
1954	4	60266	1987	37	109300
1955	5	61465	1988	38	111026
1956	6	62828	1989	39	112704
1957	7	64653	1990	40	114333
1958	8	65994	1991	41	115823
1959	9	67207	1992	42	117171
1960	10	66207	1993	43	118517
1961	11	65859	1994	44	119850
1962	12	67296	1995	45	121121
1963	13	69172	1996	46	122389
1964	14	70499	1997	47	123626
1965	15	72538	1998	48	124761
1966	16	74542	1999	49	125786
1967	17	76368	2000	50	126743
1968	18	78534	2001	51	127627
1969	19	80671	2002	52	128453
1970	20	82992	2003	53	129227
1971	21	85229	2004	54	129988
1972	22	87177	2005	55	130756
1973	23	89211	2006	56	131448
1974	24	90859	2007	57	132129
1975	25	92420	2008	58	132802
1976	26	93717	2009	59	133450
1977	27	94974	2010	60	134091
1978	28	96259	2011	61	134735
1979	29	97542	2012	62	135404
1980	30	98705	2013	63	136072
1981	31	100072	2014	64	136782
1982	32	101654	2015	65	137462
1983	33	103008	2016	66	138271

By drawing a simple scatter diagram, we can get the asymptote that the population changes nonlinearly with time and tends to be parallel to the x-axis. By drawing the scatter diagram, it can be concluded that the change of population with time was nonlinear and tends seemingly to be asymptotically parallel to the X axis. Therefore, the logistic curve model was considered. The logistic model of population forecast was as follows.

$$x(t) = \frac{x_m}{1 + \left(\frac{x_m}{x_0} - 1\right) e^{-rt}} \tag{1}$$

Where x_m is the maximum population, x_0 is the initial population in 1951, r is the natural population growth rate. Deriving the formula to get

$$\frac{x(n+1) - x(n)}{x} = r - \frac{r}{x_m} x \tag{2}$$

Let $u = \frac{x(n+1) - x(n)}{x}$, so $u = \frac{r}{x_m} x + r$, all values of u were calculated by the population data over the years.

Then, the Levenberg-Marquardt iteration method was used to calibrate the model parameters by MATLAB software, and the value of r and x_m were obtained ($r=0.036$, $x_m=1.65 \times 10^5$).

Accordingly, the logistic model of population prediction was

$$X = \frac{1.65 \times 10^5}{1 + 1.93e^{-0.036t}} \tag{3}$$

On the base of the logistic prediction model, the population change trend was shown in Figure 1. The maximum population will reach 1.65 billion in the future.

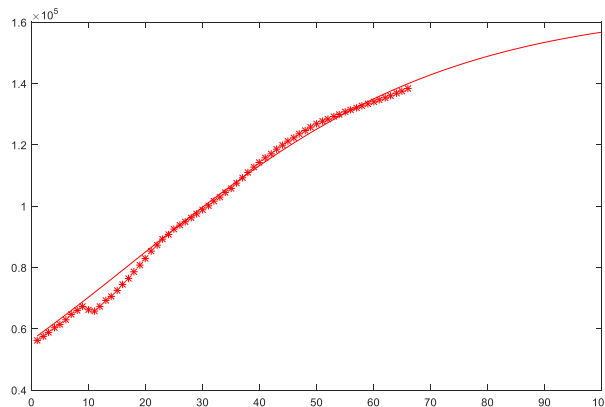


Figure 1. Demographic trends of China

This section only uses China as an example to model and predict the size of the population. For other countries or language users, please refer to this method.

III. GREY PREDICTION MODEL FOR THE NUMBER OF LANGUAGE USERS

In our globalized and interconnected world, there are many factors that allow languages that are geographically distant to interact. These factors include the language(s) promoted by the government, migration and assimilation of cultural groups, international business relations, the use of electronic communication and social media, and the use of technology to assist in quick and easy language translation [2].

Due to the above-mentioned uncertain factors, the gray prediction method is used to predict the changing trend of the number of language users. The grey prediction model was solved by MATLAB software [3].

Step 1, select the original time series data $X^{(0)}$, and make one time accumulation to generate $X^{(1)}$.

Step 2, construct the accumulation matrix B and the constant vector y .

Step 3, use the least square method to solve the gray model parameters.

Step 4, substitute the parameters of the grey model into the prediction model, and get the prediction value of the number of users of the popular languages in the prediction year.

The number of speakers of major languages in the world in recent years was shown in Table 2 and Table 3, respectively.

Table 2. The number of users of popular languages

Year	Chinese	English	Spanish	Russian	French	Hindi	Arabic	Portuguese	Bengali	German	Japanese
2008	1120.00	480.00	320.00	285.00	265.00	250.00	221.00	198.00	195.00	170.00	133.00
2009	1036.00	618.00	376.00	278.00	213.00	487.00	285.00	203.00	207.00	91.00	127.00
2010	1120.00	480.00	320.00	285.00	265.00	250.00	221.00	188.00	185.00	109.00	133.00
2011	1120.00	450.00	350.00	285.00	265.00	350.00	221.00	188.00	185.00	109.00	133.00
2012	1120.00	480.00	320.00	285.00	270.00	250.00	221.00	248.00	185.00	100.00	133.00
2013	1120.00	480.00	320.00	285.00	265.00	250.00	221.00	188.00	185.00	109.00	133.00
2014	1120.00	480.00	320.00	285.00	265.00	250.00	221.00	188.00	185.00	109.00	133.00
2015	1115.00	480.00	315.00	280.00	270.00	250.00	220.00	190.00	180.00	110.00	120.00
2016	1030.00	840.00	490.00	280.00	220.00	380.00	490.00	200.00	210.00	86.00	130.00
2017	1090.00	942.00	570.00	275.00	274.00	208.00	385.00	262.00	380.00	210.00	130.00
2018	1090.00	983.00	527.00	267.00	229.00	544.00	422.00	229.00	261.00	129.00	129.00

Table 3. The prediction value of the number of popular language users

Year	Chinese	English	Spanish	Russian	French	Hindi	Arabic	Portuguese	Bengali	German	Japanese
2019	1087.00	1216	625.1664	271.2553	252.8204	343.4786	537.62	240.2249	334.8012	170.4588	127.5426
2020	1086.00	1339.5	667.6447	270.0489	252.7167	346.4838	586.91	244.8743	355.8051	179.7083	127.2059
2021	1085.00	1475.5	713.0093	268.8478	252.613	349.5154	640.72	249.6136	378.1267	189.4597	126.8701
2022	1084.00	1625.3	761.4563	267.652	252.5093	352.5735	699.46	254.4447	401.8486	199.7402	126.5352
2023	1083.00	1790.4	813.1952	266.4616	252.4057	355.6583	763.58	259.3694	427.0587	210.5786	126.2012
2024	1082.00	1972.2	868.4496	265.2765	252.3021	358.7701	833.59	264.3893	453.8504	222.0051	125.868
2025	1081.00	2172.5	927.4583	264.0966	252.1986	361.9092	910.02	269.5063	482.3229	234.0516	125.5358
2026	1079.00	2393.1	990.4765	262.922	252.0951	365.0757	993.45	274.7224	512.5816	246.7518	125.2044

It can be seen from Figure 2 that the number of users of major languages in the world has changed many times in the past decade. According to the prediction results, English will replace Chinese as the most widely used language in the world in the future, and the number of Arabic and Spanish users was growing rapidly with a trend of surpassing Chinese.

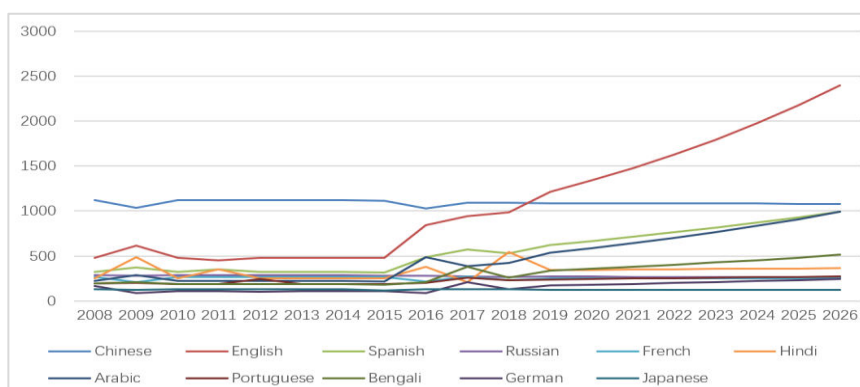


Figure 2. Trends of the number of popular language users

IV. K-MEANS CLUSTERING METHOD FOR OFFICE ADDRESS SELECT

The client company wants to set up six new international offices, the strategy of regional coverage priority was adopted. First, collected the latitude and longitude information of all cities in the world through the website [4] as shown in Figure 3.

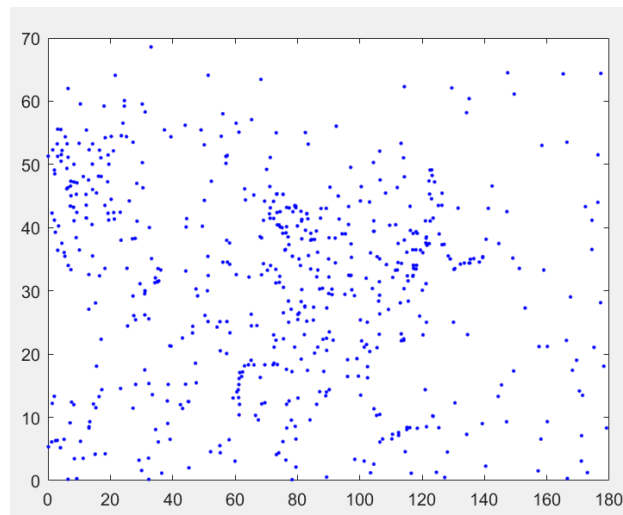


Figure 3. Location of major cities

Then used k-means clustering method to determine the specific location [5]. The algorithm steps are as follows.

Step 1: Obtain the latitude and longitude data of each major city and establish a data set.

Step 2. Randomly select 6 data points from all cities as the initial clustering center.

Step 3. Calculate the distance between each sample and each cluster center.

Step 4. Allocate each sample to the nearest cluster center. The cluster centers and the objects assigned to them represent a cluster.

Step 5. Find the centroid of the samples of each cluster as the new cluster center.

Step 6. Recalculate the distance between each sample and each cluster center. Once the samples are redistributed, the cluster center of the cluster needs to be recalculated. Until no samples are reassigned to different clusters or cluster centers no longer change.

Using MATLAB software for k-means clustering analysis, the coordinates of the six clustering centers were (15.197317, 27.042317), (24.321220162.040488), (47.273137, 14.699608) (17.713298100.707447), (33.381730, 71.579297) (39.244000, 120.506095), as shown in Figure 4.

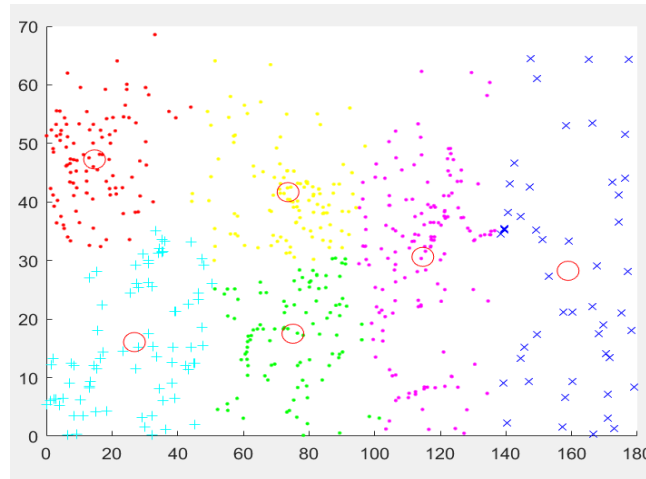


Figure 4. Clustering and clustering center

Short-term offices only consider the geographical location of the city, and each office try to cover a wider area. Therefore, the cluster center obtained by longitude and latitude clustering was used as a short-term office. The city information was shown in Table 4.

Table 4. The short-term offices location

City	Country	Continent	Language
Lusaka	Zambia	Africa	English
Noumea	New Caledonia	Oceania	French
Salzburg	Austria	Europe	German
Washington	USA	North America	English
Santiago	Chile	South America	Spanish
Beijing	China	Asia	Chinese

However, there were more factors to consider when setting up a long-term office. Not only the geographical location, but also the prosperity and comprehensive strength of the city or country should be considered, so as to promote the business and obtain better returns. Therefore, the introduction of gross domestic product (GDP) was taken into consideration.

Table 5 listed the top 20 countries with high GDP, and found that the countries with high GDP are mainly in Europe, America and Asia.

Table 5. Top 20 countries in GDP

No.	National	GDP (billion)	Continent
1	U.S.A.	\$174276	American
2	China	\$105345.3	Asian
3	Japan	\$48487.3	Asian
4	Germany	\$38967.9	European
5	Britain	\$30023.9	European
6	France	\$28538.3	European
7	Baxi	\$24560.5	American
8	Italy	\$21551.5	European
9	Russia	\$20636.6	European
10	India	\$20353.9	Asian
11	Canada	\$17928.8	American

12	Australia	\$14495.2	Oceania
13	Korea	\$14113.3	Asian
14	Spain	\$13780.4	European
15	Mexico	\$12984.7	American
16	Turkey	\$9340.75	Asian
17	Indonesia	\$8910.51	Asian
18	Netherlands	\$8810.33	European
19	Saudi Arabia	\$7563.5	Asian
20	Switzerland	\$7092.59	European

On the basis of location distribution clustering, it is suggested that Canberra should be used to replace Noumea in the location of Oceania, and Berlin should be used to replace Salzburg in the location of Europe for new international office considering the impact of GDP. The establishment of long-term international offices was shown in Table 6. Each international office was located in the capital city of the country in which it is located, considering both location coverage condition and economic development.

Table 6. The long-term offices

City	Country	Continent	Language
Lusaka	Zambia	Africa	English
Canberra	Australia	Oceania	Italian
Berlin	Germany	Europe	German
Washington	U.S.A.	North America	English
San Diego	Chile	South America	Spanish
Beijing	China	Asia	Chinese

V. CONCLUSION

The logistic curve model was adapted to predict the trend of future population change. The maximum population will reach 1.65 billion in the future. This method can provide reference for other countries' population prediction.

The grey prediction model was proposed to predict what will happen to the numbers of language speakers in the future. The result showed that the number of users of major languages in the world has changed many times in the past decade, English will replace Chinese as the most widely used language in the future, and the number of Arabic and Spanish users was growing rapidly with a trend of surpassing Chinese.

The strategy of regional coverage priority was adopted to select locations for short-term new international offices. By using MATLAB software for k-means clustering analysis, the coordinates of the six clustering centers were determined. Considering the impact of GDP, it is suggested that Canberra and Berlin should be used to replace Noumea and Salzburg respectively. Six new international offices were located in the capital city of the country considering both location coverage condition and economic development.

REFERENCES

- [1] <https://www.comap.com/undergraduate/contests/index.html>
- [2] Cao Ruo. A new study on language transfer [J]. Journal of northwestern university (philosophy and social science edition), 2012, 42(02): 158-160.
- [3] Tong Mingyu. Grey modeling method and its application in prediction [D]. Chongqing University, 2016.
- [4] <http://www.cnblogs.com/haibin-zhang/p/4955880.html>
Wang Feifei. Improved research and application of k-means clustering algorithm [D]. Lanzhou jiaotong University, 2017.
- [5] Wang Feifei. Improved research and application of k-means clustering algorithm [D]. Lanzhou jiaotong University, 2017.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details