

The Artificial Bee Colony (ABC) Algorithm for Estimating Parameter of Epidemic Influenza Model

Ririn Nirmalasari^{*}, Agus Suryanto, Syaiful Anam

Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Brawijaya, Malang, Indonesia

Abstract

The Artificial Bee Colony (ABC) is one of the stochastic algorithms that can be applied to solve many real-world optimization problems. In this paper, The ABC algorithm was used to estimate the parameter of the epidemic influenza model. This model consists of a differential system represented by variations of Susceptible (S), Exposed (E), Recovered (R), and Infected (I). The ABC processes explore the minimum value of the mean square error function in the current iteration to estimate the unknown parameters of the model. Estimating parameters were made using participation data containing influenza disease in Australia, 2017. The best parameter chosen from the ABC process matched the dynamical behavior of the influenza epidemic field data used. Graphical analysis was used to validate the model. The result shows that the ABC algorithm is efficient for estimating the parameter of the epidemic influenza model.

Keywords: ABC, Epidemic, Estimate, Influenza, Parameter.

INTRODUCTION

Parameter estimation is the process of using sample data to estimate the parameters of the optimization problem. Several parameter estimation methods are available. In this study, the chosen method is the stochastic method because of its advantages that can find the global minima solution. The determination of parameter values will affect the estimation results. If the parameter values are incorrect, the estimation results will not be accepted. It means that only the best parameter values can use so that the model can represent the field data.

The Artificial Bee Colony (ABC) algorithm is one of the most popular and widely used stochastic methods to find the solution of optimization problems. This algorithm is based on the behavior of bees to find and exploit food resources efficiently [1]. The advantages of this algorithm are simple, reliable, and flexible. The performance of the ABC algorithm is better than other stochastic algorithms such as Particle Swarm Optimization, Genetic Algorithm, and Differential Evolution on constraints or non-constraints [2].

The parameter estimation is widely studied and developed by researchers, in general, using test functions such as the Rastrigin, Ackley, Sphere, and Rosenbrock. Some researchers began to develop and apply the ABC algorithm to

estimate parameters of the model in the system of the differential equation, both linear and nonlinear. Roeva [3] used the ABC algorithm to identify the parameters of the cultivated *E. coli* bacteria model.

The aim of this study is the implementation of the ABC algorithm used to estimate the unknown parameters of the influenza epidemic model. The used mathematical model is based on the research of Samsuzzoha [4]. This model is a system of nonlinear differential equations represented by variations of Susceptible (S), Exposed (E), Infected (I), and Recovered (R) populations. The solution of this model was estimated by influenza laboratory surveillance data obtained from WHO regional databases. The chosen laboratory data is based on the well-defined standard condition.

MATERIAL AND METHOD Epidemic Influenza Model

There are four groups of population in this model: the susceptible population who can contract influenza, the exposed population who affected influenza but not to be infectious, the infected population who has infected and able to transmit the influenza virus, and the recovered population who can be re-infected by the disease if the immunity has been lost. It is assumed that the infected population only affects the susceptible population. However, among these populations, there is an exposed population that lies in a latent period or incubation period and has a low infection rate. Furthermore, the infected population that has received treatment and recovered naturally changes into a recovered population. At one time, this population could be

*Correspondence address:

Ririn Nirmalasari

E-mail : nirmala_ririn@student.ub.ac.id

Address : University of Brawijaya, Veteran Malang, 65145

re-infected by influenza when the immunity disappeared. The epidemic influenza model was expressed in the form of the following system of nonlinear differential equations [4],

$$\begin{aligned} \frac{dS}{dt} &= -\beta \frac{IS}{N} - \mu S + rN + \delta R \\ \frac{dE}{dt} &= \beta \frac{IS}{N} - (\mu + \sigma + k)E \\ \frac{dI}{dt} &= \sigma E - (\mu + \alpha + \gamma)I \\ \frac{dR}{dt} &= kE + \gamma I - \mu R - \delta R \end{aligned} \quad (1)$$

and

$$N = S + E + I + R \quad (2)$$

where the parameter description and variable model can be seen in Table 1.

Table 1. Description of all parameters and variables

Parameter	Description
β	Transmission coefficient
σ^{-1}	Incubation period
γ^{-1}	Mean duration of recovery infected population
δ^{-1}	Loss of immunity period
μ	Natural mortality rate
r	Birth rate
k	Recovery rate
α	The disease induced mortality rate
Variable	Description
S	Susceptible population
E	Exposed population
I	Infected population
R	Recovered population
N	Total population

The ABC Algorithm Implemented in Epidemic Influenza Model

In the ABC algorithm, a bee colony composes of three groups. Those are employed bee, onlooker bee, and scout bee. The number of bees in a colony is expressed as population size (*Popsiz*). Half of *Popsiz* is employed bee, and the other half is onlooker bee. There is one employed bee in each food source. It is said that the number of food sources where are around the nest (*NFood*) is equal to the number of employed bees[5].

The first step in the ABC algorithm is initialization. In this step, there are several processes: inputting parameters, setting boundary conditions, searching initial position of the food source, creating an initial population, calculating objective and fitness function, determining the initial best particle, and setting the *limit*. A food source position represents a potential parameter (x_i) as a solution to the influenza model. The initial food source position

was searched randomly by maximum (x_{max}) and minimum (x_{min}) range of parameters using equation (3). The amount of nectar in the food source determines the quality of the solution. It can be calculated by the objective function (MSE) and the fitness value, both expressed in equation (5) and (6).

In the employed bee phase, each bee modified the position of the food source randomly based on neighboring principles. It means that the chosen food source must be different from the previous food source. A new food source was generated from the previous food sources. Furthermore, these food sources are selected using a greedy selection process expressed in equation (9)–(10) to get the best food source ($pbest_i$).

In the onlooker bee phase, a bee was placed in each food source using a roulette wheel selection process, which given in equation (11). This process was based on proportional probabilities to the amount of nectar. Food sources that have the most nectar will be chosen more often by onlooker bee. Furthermore, the onlooker bee updates the food source using the same process in the employed bee phase.

Each bee colony has a scout bee that has no guidance during food search. This bee searched randomly in all directions and found a variety of food sources that have not been visited by an employed bee. It was selected by using a parameter called a limit. If a solution represents a food source that does not produce improvements in some iterations ($trial_i$), the food source discarded by the employed bee, and this bee will be a scout bee. The food sources obtained in each iteration compared with solutions in the previous iteration. If the new food source has a good quality compared to the previous food source, the new food source will be taken as the best solution ($pbest_i$). In the ABC algorithm, onlooker bee and employed bee run the exploitation process. This process applies a piece of existing knowledge to seek the best solution. On the other hand, a scout bee controls the exploration process that concerned for the entire search space of an optimal solution.

The final step is the criteria for stopping iteration. If the process has been reached maximum iteration, the process must be stopped, and the algorithm produces the best solution from all iterations ($gbest$). Based on the above explanation, the process of the ABC

algorithm implemented in the epidemic influenza model is given below [6].

Step 1: Initialization

Input parameters t (period of disease), $Popsize$ (number of population), $NFood$ (number of food source), $maxiter$ (maximum iteration), and D (dimensionality of search space);

Set the boundary condition of the parameters ($xmax$ and $xmin$);

for $i = 1, 2, \dots, NFood$ do

Position for the initial food source using equation (3);

$$x_i(0) = xmax + rand.(xmax - xmin) \quad (3)$$

where $rand$ is a random number [0,1];

Create an initial population using the SEIRS model (equation (1)-(5)) and the Runge-Kutta 4th order method given as below;

$$\begin{aligned} k_1 &= h.f(t_i, x_i) \\ k_2 &= h.f(t_i + \frac{h}{2}, x_i + \frac{k_1}{2}) \\ k_3 &= h.f(t_i + \frac{h}{2}, x_i + \frac{k_2}{2}) \\ k_4 &= h.f(t_i + h, x_i + k_3) \\ x_{i+1} &= x_i + \frac{1}{6} \cdot (k_1 + 2k_2 + 2k_3 + k_4) \end{aligned} \quad (4)$$

where $f(t_i, x_i)$ is model of influenza, h is interval of time, and x_{i+1} is the solution of model;

Calculate objective functions using MSE;

$$f_i = \frac{1}{t} \sum_{i=1}^t (data_i - x_i)^2 \quad (5)$$

Calculate fitness values using equation (6);

$$fit_i = \begin{cases} \frac{1}{1+f_i}, & jikaf_i > 0 \\ 1 + abs(f_i), & jikaf_i < 0 \end{cases} \quad (6)$$

end for;

Determine the initial $pbest$ and $gbest$;

Set iteration = 0;

Set $trial = 0$;

Set $limit = NFood \times D$; (7)

Step 2: The employed bee phase

for $i = 1, 2, \dots, NFood$ do

Update new food sources using equation (8);

$$X_i(iteration + 1) = x_i(iteration) + \psi \cdot (x_i(iteration) - x_k(iteration)) \quad (8)$$

where $k = 1, 2, \dots, NFood$ and ψ is random number [-1,1];

Determine the position of food sources to stay within the range;

Generate new populations using the SEIRS model (equation (1)) and the Runge-Kutta 4th order method (equation (4));

Calculate objective functions using MSE;

Determine fitness values using equation (6);

Select the best food source between x_i and X_i using the Greedy process below,

If $fit_{i+1} > fit_i$, then

$$x_i(iteration + 1) = X_i(iteration + 1) \quad (9)$$

$$trial_i = 0$$

else

$$trial_i = trial_i + 1 \quad (10)$$

end if

end for

Step 3: Calculate P_i Opportunities using equation (11)

$$P_i = \frac{fit_i}{\sum_{j=1}^{NFood} fit_j} \quad (11)$$

Step 4: The onlooker bee phase

for $i = 1, 2, \dots, NFood$ do

if $rand(0,1) < P_i$, then

Do Employed bee phase

end if

end for

Step 5: The scout bee phase

if $max(trial_i) < limit$, then

Replace x_i with the new value using equation (3);

Generate x_i population using the SEIRS model (equation (1)) and the Runge-Kutta 4th order method;

Calculate objective functions using MSE;

Determine fitness values using equation (6)

end if

If x_i is better than $pbest_i$, then

Set x_i to $pbest_i$

end if

Determine the particle with the smallest MSE to be the best particle;

$$iteration = iteration + 1 \quad (12)$$

Step 6: If the iteration reached the maximum iteration, the process was stopped. Instead, the process should be repeated in step 2.

RESULT AND DISCUSSION

Setting of experiment

The ABC algorithm was coded in Matlab R2013a, and the experiment was implemented 10 times with a different number of populations. It is applied to estimate the parameter of the epidemic influenza model based on infected influenza data in Australia, 2017. This data was presented in Table 3.

The maximum iteration was set to 1000 iterations. The initial condition of parameters and populations showed in Table 2.

Table 2. The initial condition of parameters and populations

Parameter	Value	Source
<i>Popsize</i>	25, 50, 75, 100	Assumption
t	52 (week)	[7]
Dimension (D)	4	Assumption
N	24585 (thousand)	[8]
r	0.0248 (per 100 population per week)	[8]
μ	0.01269 (per 100 population per week)	[8]
σ	[1.75, 7] (per week)	[9]
γ	[1, 1.4] (per week)	[9]
α	0.000098 (per 100 population per week)	[10]
δ	0.0192 (per week)	[4]
β	[4.06e-6, 50] (per 100 population per week)	Assumption
k	[4.06e-6, 50] (per 100 population per week)	Assumption
S_0	24584.372 (thousand)	Assumption
E_0	0.557 (thousand)	Assumption
I_0	0.036 (thousand)	[8]
R_0	0.035 (thousand)	Assumption

The model of influenza has eight parameters. There are four fixed parameters, namely r , μ , α , and δ , based on Table 2. On the other hand, four unknown parameters, which are σ , γ , β , and k , will be estimated using the ABC algorithm. The initial condition of populations was determined based on Australia Influenza Laboratory Surveillance Information in the first week and several other assumptions.

Model validation

Data in Table 3 was simulated using the ABC algorithm, and the estimation results can be seen in Figure 1 and Figure 2. Figure 1 shows the ability of the ABC algorithm to find global minima solutions. It has been seen that during the searching process until the maximum iteration, the ABC algorithm search for the smallest MSE value of each iteration. The ABC algorithm produces a different solution for each experiment. It due to the selection of population or prospective solution was selected randomly. The convergence of the ABC algorithm was faster when the population sizes were large. It can be seen in Table 4, which smallest fitness value lays in the fifth experiment in population size of 100.

The best parameter value that corresponds to the best fitness values of each *Popsize* was entered in equation (1), and the results can be seen in Figure 2. Most of the simulation results close to field data values, although some data are difficult to reach by the ABC algorithm. The standard deviation value of each experiment with different *Popsize* is always less than the mean of the best solution and decreases with increasing *Popsize*. It means that the estimation data was not too diverse, even though the standard deviation value is high due to the high value of the infected population. However, overall the estimated parameters of the ABC algorithm can represent well the model of the spread of influenza disease, especially in Australia.

Finally, the best parameters from all experiments are $\beta = 2.23$, $\sigma = 5.63$, $\gamma = 1.32$, and $k = 3.12$ (see in Table 4). β parameter means that 2.23% of population will be affected by influenza in a week. Affected people are in an average incubation period for a day. During this period, 3.12% of the population heals from influenza, while people who don't recover become infected populations. This population heals from influenza after those take medicine and medical treatment for five days.

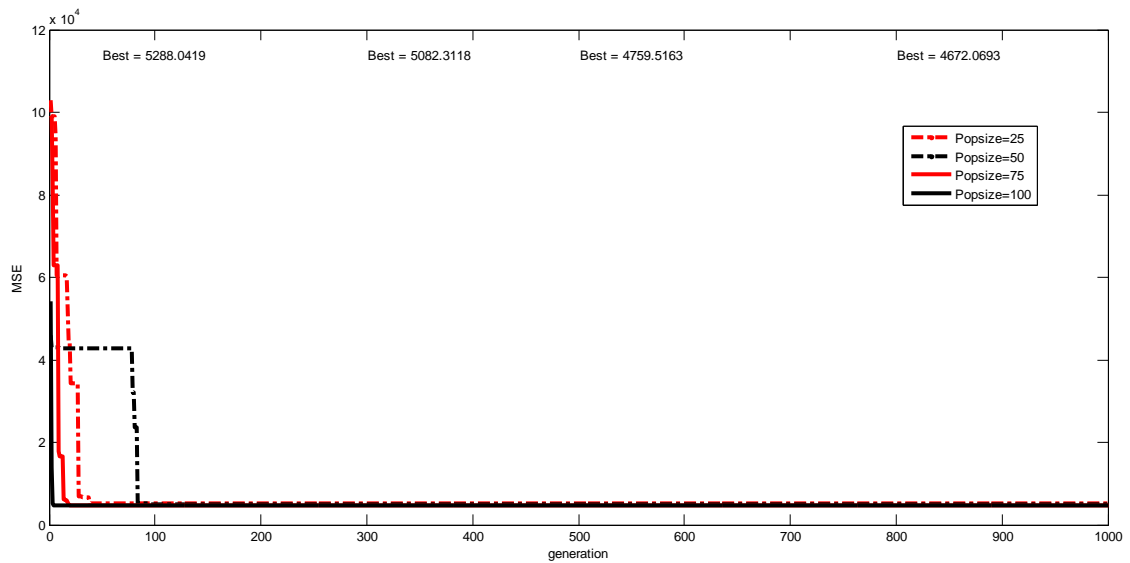


Figure 1. Convergence curves of ABC algorithm

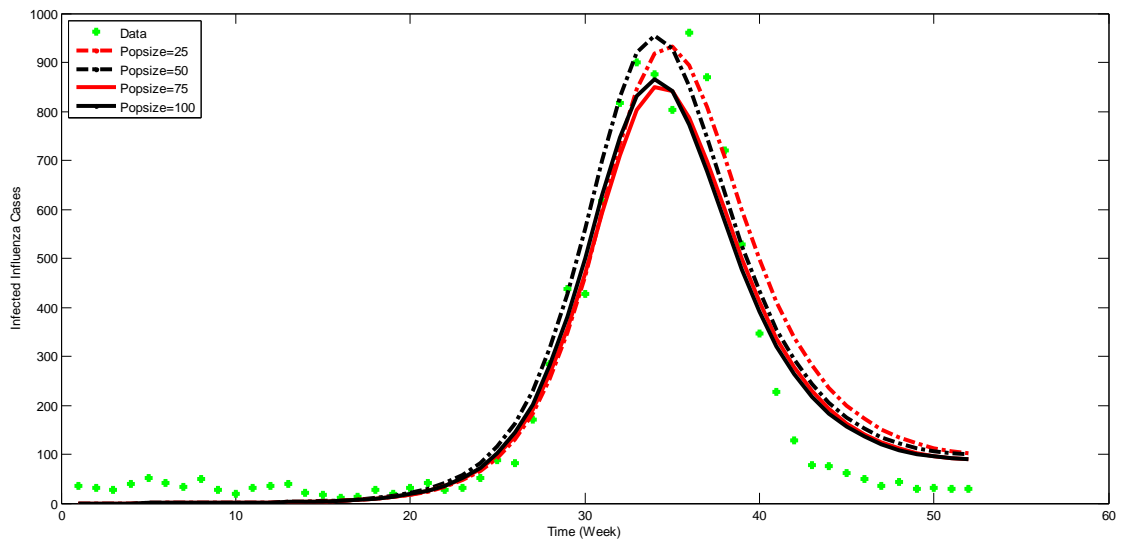


Figure 2. Comparison of field data and estimation results based on parameter values that correspond to the best fitness values for each *Popsiz* in Table 4

CONCLUSION

The result shows that the ABC algorithm is efficient for estimating the unknown parameter of the epidemic influenza model. It is due to most of the simulation results are close to field data values. It means that the simulation result represents well the real condition of the epidemic influenza. The population size is influential in obtaining more improved for estimated parameters. Our future study will be focused on the implementation of an improved ABC algorithm or another stochastic method on the dynamical model.

REFERENCES

- [1] Karaboga, D. 2005. An idea based on honey bee swarm for numerical optimization. Technical Report-Tr06. Turki.
- [2] Karaboga, D., B. Akay. 2009. A comparative study of Artificial Bee Colony algorithm. *Appl. Math. Comput.* 214(1). 108-132. Doi: 10.1016/j.amc.2009.03.090.
- [3] Roeva, O. 2018. Application of artificial bee colony algorithm for model parameter identification. *Innovative Computing: Optimization and Its Applications, Studies in Computational Intelligence.* 741. DOI: 10.1007/978-3-319-66984-7_17.
- [4] Samsuzzoha, M., M. Singh, D. Lucy. 2011. Numerical study of diffusive epidemic model of influenza with variable transmission coefficient. *Appl. Math. Model.* 35. 5507-5523.
- [5] Suyanto. 2017. Particle swarm optimization. *Swarm Intelligence, Modern Computation*

- for Optimization Big Data Mining. Chapter 4. 51-91.
- [6] Li, Z., W. Wang, Y. Yan, Z. Li. 2015. PS-ABC: a hybrid algorithm based on particle swarm and Artificial Bee Colony for high-dimensional optimization problems. *Expert Syst. Appl.* 42(22). 8881-8895. DOI: 10.1016/j.eswa.2015.07.043.
- [7] WHO. Influenza: FluNet. 2019. Available at: https://www.who.int/influenza/gisrs_laboratory/flunet/en/.
- [8] United Nations, Department of Economic and Social Affairs, Population Division. 2019. World Population Prospects 2019, Volume I: Comprehensive Tables (ST/ESA/SER.A/426). United Nations
- [9] WHO. Influenza (seasonal). 2019. Available at: [https://www.who.int/en/news-room/fact-sheets/detail/influenza-\(seasonal\)](https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal)).
- [10] Australian Bureau of Statistics. 2017. Australia's leading causes of death, 2017. Available at: <https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/3303.0~2017~Main%20Features~Australia's%20leading%20causes%20of%20death,%202017~2>.

Table 3. Influenza Laboratory Surveillance Information in Australia, 2017

Year	Start date	End date	Total number of influenza positive viruses	Start date	End date	Total number of influenza positive viruses
2017	2017-01-02	2017-01-08	36	2017-07-03	2017-07-09	170
2017	2017-01-09	2017-01-15	32	2017-07-10	2017-07-16	284
2017	2017-01-16	2017-01-22	28	2017-07-17	2017-07-23	437
2017	2017-01-23	2017-01-29	39	2017-07-24	2017-07-30	428
2017	2017-01-30	2017-02-05	51	2017-07-31	2017-08-06	619
2017	2017-02-06	2017-02-12	41	2017-08-07	2017-08-13	817
2017	2017-02-13	2017-02-19	33	2017-08-14	2017-08-20	901
2017	2017-02-20	2017-02-26	50	2017-08-21	2017-08-27	876
2017	2017-02-27	2017-03-05	27	2017-08-28	2017-09-03	804
2017	2017-03-06	2017-03-12	20	2017-09-04	2017-09-10	960
2017	2017-03-13	2017-03-19	32	2017-09-11	2017-09-17	869
2017	2017-03-20	2017-03-26	36	2017-09-18	2017-09-24	721
2017	2017-03-27	2017-04-02	40	2017-09-25	2017-10-01	529
2017	2017-04-03	2017-04-09	21	2017-10-02	2017-10-08	347
2017	2017-04-10	2017-04-16	18	2017-10-09	2017-10-15	228
2017	2017-04-17	2017-04-23	11	2017-10-16	2017-10-22	128
2017	2017-04-24	2017-04-30	14	2017-10-23	2017-10-29	78
2017	2017-05-01	2017-05-07	27	2017-10-30	2017-11-05	76
2017	2017-05-08	2017-05-14	19	2017-11-06	2017-11-12	62
2017	2017-05-15	2017-05-21	32	2017-11-13	2017-11-19	49
2017	2017-05-22	2017-05-28	41	2017-11-20	2017-11-26	35
2017	2017-05-29	2017-06-04	27	2017-11-27	2017-12-03	43
2017	2017-06-05	2017-06-11	31	2017-12-04	2017-12-10	30
2017	2017-06-12	2017-06-18	51	2017-12-11	2017-12-17	32
2017	2017-06-19	2017-06-25	88	2017-12-18	2017-12-24	30
2017	2017-06-26	2017-07-02	82	2017-12-25	2017-12-31	29

Data source: FluNet (www.who.int/flunet), Global Influenza Surveillance and Response System (GISRS) [7].

Table 4. The Performance of ABC Algorithm

Popsiz	Test	Fitness Value	Best	Worst	Mean	Standard Deviation	Parameter Value			
							β	σ	γ	k
25	I	6152.66					2.69	3.19	1	4.19
	II	6288.43					2.59	1.75	1.29	1.29
	III	9377.17					2.91	7	1	10.88
	IV	10317.4					2.85	7	1.06	9.92
	V	9935.48					3.00	5.20	1	8.43
	VI	5288.04	5288.04	10915.2	7835.29	2073.42	2.05	7	1.23	3.72
	VII	6317.16					2.49	7	1.12	6.99
	VIII	7337.14					2.02	6.86	1.13	4.18
	IX	6424.23					2.65	3.43	1.13	3.63
	X	10915.2					2.65	5.06	1.4	3.82
50	I	5082.31					2.07	5.57	1.27	2.69
	II	8991.52					2.19	3.95	1.00	3.47
	III	6166.98					2.22	5.48	1.07	4.57
	IV	9040.99					2.85	7	1.03	10.16
	V	6166.98					2.22	5.48	1.07	4.57
	VI	10930.9	5082.31	10930.9	7535.53	2070.4	2.10	3.50	1.05	2.53
	VII	5082.31					2.07	5.57	1.27	2.69
	VIII	8991.52					2.19	3.95	1.00	3.47
	IX	8991.52					2.19	3.95	1.00	3.47
	X	5910.21					2.53	6.17	1	7.45
75	I	4759.52					2.39	3.24	1.32	2.04
	II	5784.37					2.42	6.89	1	7.71
	III	5263.1					2.35	5.97	1.21	4.55
	IV	4759.52					2.39	3.24	1.32	2.04
	V	5366.74					2.27	6.87	1.09	5.86
	VI	5263.1	4759.52	10140.6	5727.19	1602.13	2.35	5.97	1.21	4.55
	VII	5912.36					2,48	7	1	8.19
	VIII	4759.52					2.39	3.24	1,32	2.04
	IX	10140.6					2.97	7	1	11.26
	X	5263.1					2,35	5.97	1.21	4.55
100	I	6042.76					2.52	6.81	1.06	7.61
	II	5517.69					2.39	6.35	1.05	6.45
	III	5176.03					2.25	3.56	1,20	2.34
	IV	7085.33					2,12	4.55	1.12	3.07
	V	4672.07	4672.07	7085.33	5913.48	853.29	2.23	5.63	1.32	3.12
	VI	5300.55					2.31	4.79	1.11	4.06
	VII	6940.62					2.67	5.87	1.04	7.40
	VIII	5300.55					2.31	4.79	1.11	4.06
	IX	6940.62					2.67	5.87	1.04	7.40
	X	6158.6					2.71	3.51	1	4.66