Approach of Implementation using models of analysis of main component for breeding and reproduction of tilapia

Sergio Barrios[1], Alberto Ochoa[2]

Abstract—Nowadays, aquaculture projects are beginning to be carried out in many parts of Mexico, where the main obstacle to its success is the lack of knowledge of the basic principles and technical necessary skills to achieve maximum utilization. In Mexico, the projection of a tilapia hatchery has required models established in other countries, making the necessary modifications according to the natural conditions of the country. This work intends to apply multivariate analysis techniques and especially based on main components, due to it is a method widely used to reduce the size of a group of specific data, locating it in the specific variables according to the necessity that is required, maintaining a High-level order of the main components which are determined and ignoring the low level for the decision making. They optimize according to the natural conditions of Mexico and especially the state of Querétaro in the breeding and reproduction of tilapia, constructing predictive models, at the time of developing applicable projects. For this, the first section of this chapter is a review of different concepts related to methods that may be applicable in decision making for the particularity of said project explaining methods of similar solutions giving different points of view for the solution of the same, in later sections will apply numerical methods of analysis that work with the data generated from the analysis based on main components giving a solution of the problem showing concrete results as desired. Finally, it is presented conclusions and references consulted for the accomplishment of this investigation.

Index Terms—Tilapia reproduction; implementation; models of analysis.

1 INTRODUCTION

In recent years, the global aquaculture production has increased at a sufficiently high rate to compensate for declining catches in fisheries, so aquaculture production in many countries has contributed greatly to the annual production of fishery products at the level national. The current situation of world aquaculture is characterized by the progress in the cultivation of shrimp and marine fish; and in several countries the economic income of aquaculture as a result of the export of its products represents the most important source of foreign exchange in its economy. Aquaculture projects are beginning to take place in many places, where the main obstacle to their success is the lack of knowledge of the basic principles and the necessary technical skills. In Mexico, the projection of an aquaculture farm requires models established in other countries, making the necessary modifications according to the natural conditions in the country. The development of aquaculture in Mexico compared to the world aquaculture development shows a lag in both the diversity and use of resources as in the modernization of the sector. Mexico is in an ideal situation for the cultivation of several species because it has extensive coastlines, abundant wetlands and optimal climates and fish farming has become a necessity in many places where this activity was not previously practiced [8]. Investigations were carried out under commercial conditions in an intensive tilapia farm in the municipality of Alvarado, Veracruz where Fifteen cages were installed inside a rustic pond of 0.5 ha, and each cage was planted with 7 000 offs of 6 g average weight yielded results Positive in terms of profits generated, Querétaro as a state, has characteristics similar to that of the municipality Alvarado, currently in Querétaro through the media, have been given more publicity and is quoted in the market between 35 mxn and 45 mxn per kilogram [6]. This is a very good alternative to improve the economic conditions of local fishing communities, directly benefiting families by generating permanent jobs, offering a better quality of life, producing low cost protein food as well as being a production alternative in the fishing activity. It is worth mentioning that the reproduction of the tilapia in the fattening phase can lead to the collapse of the cultivation system, due to the large number of tilapias that are born in the ponds and due to this the tilapias in cultivation present a low growth and a Bad conversion of food, due to the process of reproduction in which great energy is invested, a good option is to fatten lots of masculinized tilapias (The male grows faster than the female) which is achieved by supplying masculinizing hormones in the early stages when given Sexual differentiation. Although the most common practice is to keep males and females together indeterminately by separating with the nets the small fry that are born in the pond, it is best that the incubation be carried out in a controlled manner. One way to achieve this is by milking the fertilized eggs from the female's mouth, which incubates them until they are born, the eggs are incubated in clean, oxygenated water to prevent the onset of disease, the ad

- [1]*Universidad Autónoma de Querétaro, Facultad de Informática Av. de las Ciencias S/ N, Campus Juriquilla, Querétaro, Qro. C.P. 76230 México.*
- [2]*Universidad Autónoma de Ciudad Juárez, Maestria en Cómputo Aplicado, Ciudad Juárez, Chih. E-mail: alberto.ochoa@uacj.mx.*

vantage of this system is that tilapia Are fed with hormones from the time they hatch and can eat what results in a good% of male tilapia, this process can take place approximately one month after spawning. So, in this case will work with samples of masculinized issues fry for the maximum profit and profit of the business. It is worth emphasizing that when the stocking density of the tilapia is low, it is possible to dispense with oxygen aerators and only to carry out water refills, in large areas, the low intensity crop can shed interesting production volumes that are the importance to perform a good study to achieve maximum efficiency in breeding and reproduction of tilapia. Apply several techniques to apply when getting the product present quality when it reaches its final destination that is marketing for it tilapias must have complete fins without fraying, no white spots or ulcers or scratching injuries behavior against the wall or the background [5] [8].

Nowadays, there are not mathematical models to optimize resources of scientific projects, so the present research is a milestone in the subject seeking a mathematical solution. Additionally, although it will be tested on tilapia breeding projects already implemented, the research will serve as a basis for any similar problem that requires cost optimization. Finally, it will serve to focus the use of economic resources in a more real way. The main social contribution of the research will be to expand the knowledge of the subject in this branch that serves to the future investors. It will be achieved as the budget of each project is determined in real form, to avoid overcharging and underestimating them.

The main objective of this work is to determine the best implementation of a breeding and reproduction project of tilapia depending on the budget to be invested applying mathematical algorithms that help the decision making in carrying out the project.

1.3 Initial conditions.

For the implementation of this project it is worth saying that the investor must take into account that the budget that wishes to invest will have to consist of a certain limit since it is necessary to invest in minimum tools indispensable for the control and monitoring of the process, in addition to taking into account the distance in which our source of supply will be found to determine the cost of the piping system to be applied. The study is based on tanks of geomembrane supplied by professional companies in the sector and it will take into account the different capacities of the same that are tanks of (3, 6, 9, 12) m wide and 1.20 m high because they are the most used in This type of projects applied to the creation of tilapia (Extensive, Intensive, Semi-Intensive, Super-Intensive) [5].

Table 1. Elements of the Mathematical Model. It is worth to say that it is more feasible to use a methodology of planting of the crop according to the number of ponds that are implemented for the maximum use of the space of the ponds for example.

Description1	Description
Cp	Number of fish for the crop
Pi	Initial budget
PR	Actual budget (suggested by the algorithm)
Com	Cost per square meter
Cmd	Number of square meters available
CmR	Actual square meters (suggested by the algorithm)
IE	Estimated cost of infrastructure and equipment requested for cultivation
Cop	Cost of each Alevin (0.5 g) (Constant Value mxn 0.75)
DmT	Diameter of the pond to be used (3,6,9,12 meters of Diameter)
CoAl	Generalized feeding cost per individual (Constant Value 1.7 kg - 12.50 mxn / kg)
EBlo	Existence of blower in the project (1 (SI), 0 (NO))
Cpm	Number of fish that occupies one cubic meter (Constant Value 50-100 individuals)
CAg	Quantity of general food
CantT	Number of ponds to be used according to diameter required
M^2T	Metro Square according to pond (Area of pond = 3.14 * r^2)
MtCult	Cultivation Methodology to use

2 METHODOLOGY

2.1 Methodology implemented

When the number of pond is equal to or greater than 7, the pond capacity is maximized at the time of planting of the fry as it can plant the pods in full according to the construction follows:

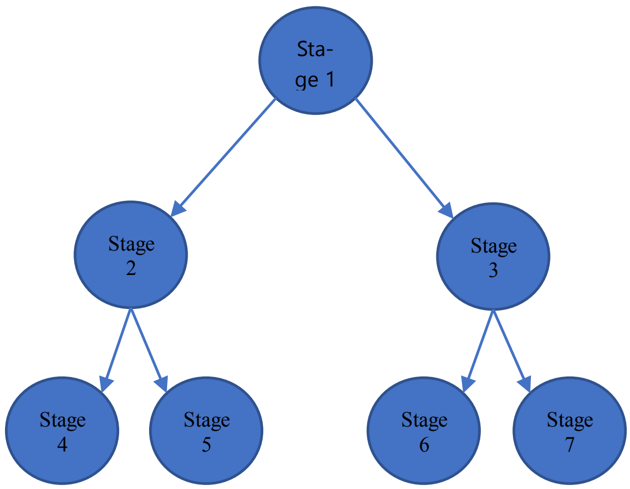

Fig. 1. Pond assignations (Source: own elaboration)

In which the capacity of the pond is fully exploited, respecting the planting rate per cubic meter (100 u / m^3)

where at the end of the breeding stage the fish reach a weight of 150 g / u (Pond 1) Dividing the population of individuals in the following ponds (Pond 2 and Pond 3), thus respecting the suitability of individual space, and in turn, at the end of the pre-fattening stage (300 g / Other ponds (Pond 4, Pond 5, Pond 6, Pond 7) until reaching the weight per individual of 500 g / u which would be the best marketing weight [10].

2.2 Methodology 2

When the number of ponds is less than 7 does not comply with the methodology above exposed and as you cannot divide populations equally as the previous method does not maximize the capacity of the pond at the time of planting because you have to think in terms of Commercialization so that it is necessary to plant the optimal capacity by square meters that support individuals of 500 g / u in the pond so the planting that is carried out in the pond will not cover the whole of the pond space since the fry would be sown to Lower amount per cubic meter of water and should be adequately concentrated. A UML modeling that explains such a situation is as follows [10]:

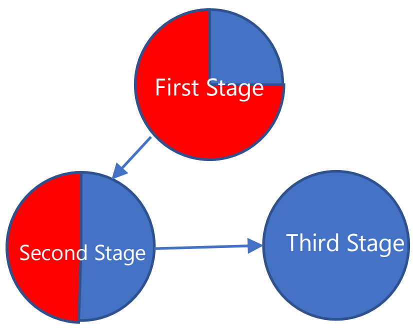

Fig. 2. UML Modeling (Source: Saavedra, 2013)

Function Purpose.

$$z = \min(\sum_{i=1}^{n} C_{om_i} + IE + \mathrm{CAg}) \quad (1)$$

Model Restrictions

The actual budget cannot exceed the initial budget.

$$P_i \leq P_r \quad (2)$$

The number of square meters used cannot exceed those available.

$$C_{mr} \leq C_{md} \quad (3)$$

Where

Quantity of ponds to be used according to Cmd.

$CantT \simeq \mathrm{Cmd}/3,\ 14*r^2$ Calculation of the number of ponds per m^2 available.

$C_{mr} = CantT * 3.14 * r^2$ Referes to each meter used with respect to the diameter of the pond used.

Estimation of the architectural method to be implemented.

$\mathrm{MtCult}^1 = CantT/7$ Amount of times I can apply methodology 1 according to ponds.

$\mathrm{MtCult}^2 = Mod(CantT/7)$ Amount of times I can applymethodology 2 according to ponds.

MtCult = MtCult^1 y MtCult^2 Number of times each methodology can be applied.

Mathematical model for methodology 1

Amount of fish per cubic meter:

$$C_{pm} = 100\ individuos$$

Cost of alevin per cubic meter general:

$$C_{om} = C_{pm} * C_{op} \quad (4)$$

The general cost of planting the alevin.

$$\mathrm{Cp} = CantT * M^3T * C_{pm} \quad (5)$$

Amount of feed to be used throughout the breeding cycle up to 500 g per fish.

$$CAg = Cp.* \ CoAl \quad (6)$$

The estimated cost for infrastructure and equipment is estimated.

$$IE = \sum_{i=1}^{n} C_{oe_i} + \sum_{j=1}^{m} C_{oi_j} \quad (7)$$

The cost of investment of the fry and feed will comprise between 50% and 60% of the initial budget.

$$P_i * 0.5 \leq \mathrm{CAg} + \mathrm{Cp} \leq P_i * 0.6 \quad (8)$$

The cost of infrastructure and equipment will comprise between 40% and 50% of the initial budget.

$$P_i * 0.4 \leq IE \leq P_i * 0.5 \quad (9)$$

Mathematical model for methodology 2

Amount of fish per cubic meter:

$$C_{pm} = 50\ individuos$$

Cost of alevin per cubic meter general:

$$C_{om} = C_{pm} * C_{op} \quad (10)$$

The general cost of planting the alevin.

$$\mathrm{Cp} = CantT * M^3T * C_{pm} \quad (11)$$

Amount of feed to be used throughout the breeding cycle up to 500 g per fish.

$$CAg = Cp.* \ CoAl \quad (12)$$

The estimated cost for infrastructure and equipment is estimated.

$$IE = \sum_{i=1}^{n} C_{oe_i} + \sum_{j=1}^{m} C_{oi_j} \quad (13)$$

The cost of investment of the fry and feed will comprise between 50% and 60% of the initial budget.

$$P_i * 0.5 \leq \mathrm{CAg} + \mathrm{Cp} \leq P_i * 0.6 \quad (14)$$

The cost of infrastructure and equipment will comprise between 40% and 50% of the initial budget.

$$P_i * 0.4 \leq IE \leq P_i * 0.5 \quad (15)$$

2.3 Analysis Based on Main Components

The so-called Common Principal Component Analysis (ACP) is where you collect information from a sample of data, most often taking as many variables as possible. However, if we take too many variables on a set of objects, for example 20 variables, we will have to consider 180 possible correlation coefficients; If there are 40 variables, that number increases to 780. Obviously, in this case it is difficult to visualize relationships between variables. Another problem that arises is the strong correlation that often occurs between variables: if we take too many variables (which usually happens when you do not know too much about the data or only have an exploratory spirit), it is normal that they are related or that they measure the same from different points of view. For example, in medical studies, blood pressure at the exit of the heart and at the exit of the lungs is strongly related. It is necessary, therefore, to reduce the number of variables. It is important to highlight the fact that the concept of greater information is related to the one of greater variability or variance. The greater the variability of the data (variance), the more information is considered, which is related to the concept of entropy [2].

2.4 Principal Components

These techniques were initially developed by Pearson at the end of the 19th century and were later studied by Hoteling in the 1930. However, the appearance of computers did not begin to popularize. In order to study the relationships that exist between correlated variables (that measure common information), the original set of variables can be transformed into another set of new variables that are not correlated with each other (which does not have a redundancy or redundancy in the information). The new variables are linear combinations of the previous ones and are constructed according to the order of importance in terms of the total variability they collect from the sample. Ideally, we look for m<p variables that are linear combinations of the original p and are uncorrelated, collecting most of the information or variability of the data. If the original variables are incorrectly matched, then it is meaningless to perform a principal component analysis. The main components analysis is a mathematical technique that does not require the multivariate normality assumption of the data, although if the latter is met it can be given a deeper interpretation of these components, it is applied, in this investigation to control the quantity and capacity of ponds to be implemented with the aim of maximizing planting space [1].

2.5 Analysis based on principal components

The so-called Common Principal Component Analysis (ACP) is where you collect information from a sample of data, most often taking as many variables as possible. However, if we take too many variables on a set of objects, for example 20 Variables, we will have to consider 180 possible correlation coefficients; If there are 40 variables, that number increases to 780. Obviously, in this case it is difficult to visualize relationships between variables. Another problem that arises is the strong correlation that often occurs between variables: if we take too many variables (which usually happens when you do not know too much about the data or only have an exploratory spirit), it is normal that they are related or that they measure the same from different points of view. For example, in medical studies, blood pressure at the exit of the heart and at the exit of the lungs is strongly related. It is necessary, therefore, to reduce the number of variables. It is important to highlight the fact that the concept of greater information is related to the one of greater variability or variance. The greater the variability of the data (variance), the more information is considered, which is related to the concept of entropy [2].

3 EXPERIMENTATION

3.1 Calculation of main components

It is considered a series of variables (x^1, x^2... xp) on a group of objects or individuals and try to calculate from them a new set of variables y^1, y^2... yp, between each other, whose variances are decreasing progressively. Each yj (where j = 1... p) is a linear combination of the original x^1, x^2... xp, that is:

$$Yj = aj^1x^1 + aj^2x^2 + \ldots + ajpxp = a0jx$$

Where $a^oj = (a^1j, a^2j \ldots apj)$ is a vector of constants, and

Obviously, if we want to maximize the variance, as we shall see later, a simple way might be to increase the coefficients aij. Therefore, in order to maintain the orthogonally of the transformation it is necessary that the vector module.

$A0j = (a^1j, a^2j \ldots apj)$ is

1. That is,

$$a_j' a_j = \sum_{k=1}^{p} a_{kj}^2 = 1 \qquad \mathbf{(16)}$$

The first component is calculated by choosing a^1 so that y^1 has the largest possible variance. The second main component is calculated by obtaining a^2 so that the obtained variable, y^2 is wrong with y^1. In the same way, y^1, y^2, and p, are wrongly chosen so that the random variables obtained will have less and less variance [1].

Table 2. Variables of study case.

Cases	Dtanque	CantT	M^3 Tanque	Palevin	Paliment	Ptanque	Can-tAlev	Pin-version	Eblo	M^2Espacio
Case A	3	2	20	750	12500	8200	1000	24160	0	28
Case B	3	3	29	1088	18125	12300	1450	33885	0	37
Case C	3	4	38	1425	23750	16400	1900	57610	1	46
Case D	3	6	56	2100	35000	24600	2800	77060	1	64

Cas				487	812	287	650	127	1	
e E	3	7	65	5	50	00	0	410		73
Cas				127	212	210	170	457	0	
e F	6	2	34	5	50	00	0	10		82
Cas				187	312	315	250	662	0	
e G	6	3	50	5	50	00	0	10		118
Cas				247	412	420	330	100	1	
e H	6	4	66	5	50	00	0	710		154
Cas				367	612	630	490	141	1	
e I	6	6	98	5	50	00	0	710		226
Cas			11	855	142	735	114	233	1	
e J	6	7	4	0	500	00	00	460		262
Cas				270	450	404	360	888	0	
e K	9	2	72	0	00	00	0	60		172
Cas			10	401	668	606	535	130	0	
e L	9	3	7	3	75	00	0	935		253
Cas			14	532	887	808	710	187	1	
e M	9	4	2	5	50	00	0	010		334
Cas			21	795	132	121	106	271	1	
e N	9	6	2	0	500	200	00	160		496
Cas			24	185	308	141	247	467	1	
e O	9	7	7	25	750	400	00	610		577
Cas			15	585	975	612	780	162	0	
e P	12	2	6	0	00	00	0	160		298
Cas			23	873	145	918	116	240	0	
e Q	12	3	3	8	625	00	50	885		442
Cas			31	116	193	122	155	333	1	
e R	12	4	0	25	750	400	00	610		586
Cas			46	174	290	183	232	491	1	
e S	12	6	4	00	000	600	00	060		874
Cas			54	405	676	214	541	907	1	101
e T	12	7	1	75	250	200	00	910		8

Where:
Dtanque Tank Diameter
CantT Amount of Tank
M^3Tanque Tank M^3
Palevin Price of alevin x Amount of Alevin
Paliment Price of feed x quantity of Alevin
Ptanque Price of the pond with its aggregates
CantAlev Number of fry x pond
Pinversion Total, Investment Used
M^2Espacio Space in M^2 used in each case

Table 3. Costs of materials.

Tools	Price
Oximeter	1400
Thermometer	160
L. de amonio	500
Motobomba	1400
Price alevin	0,75
M^3Tanque 3	9
M^3Tanque 6	16
M^3Tanque 9	35
M^3Tanque 12	77
Blower	14000
Price T 3	4100
Price T 6	10500
Price T 9	20200
Price T 12	30600

3.2 Principal components analysis

Table 4. Own values

	F1	F2	F3	F4	F5	F6	F7	F8
				0,	0,			
	7,			0	0			
Own	48	1,0	0,3	7	1	0,0		
value	9	62	53	7	9	00	0,000	0,000
				0,	0,			
Vari-	83			8	2			
ability	,2	11,	3,9	5	1	0,0		
(%)	12	798	22	0	5	03	0,000	0,000
						10		
						0,	10	10
Accumulated	83,	95,0	98,9	99,	99,9	00	0,0	0,0
(%)	212	10	32	782	97	0	00	00

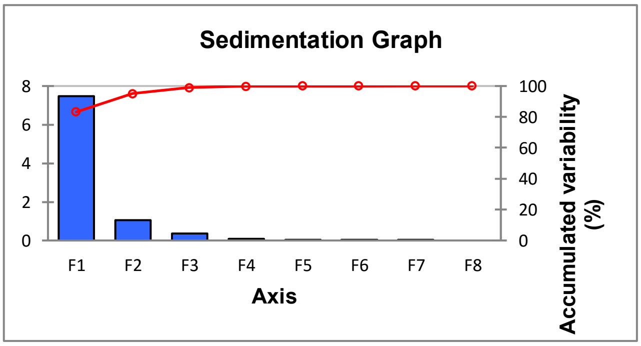

Fig. 3. Sedimentation (Source: own elaboration)

From where we can infer in the matrix of own values and in the sedimentation graph that the first 2 components have a high value in terms of% concentration of model information.

Table 5. Measurement of Kaiser-Meyer-Olkin Sample Adequacy.

Variables	Factor
Dtanque	0,979
CantT	0,391
M^3Tanque	0,665
Palevin	0,718
Paliment	0,718

Ptanque	0,664
CantAlev	0,718
Pinversion	0,880
M^2Espacio	0,667
KMO	0,704

Also, giving acceptable results in the sample referenced utilize Kaiser-Meyer-Olkin method.

Table 6. Contribution of variables.

Variables	F1	F2
Dtanque	6,961	37,482
CantT	4,005	55,157
M^3Tanque	12,637	1,549
Palevin	12,598	0,953
Paliment	12,598	0,953
Ptanque	12,730	0,420
CantAlev	12,598	0,953
Pinversion	13,152	0,540
M^2Espacio	12,720	1,992

Table 7. Square cosines of observations. It represents the % of belonging of each case assigning it to the component of greater probability [10].

Observations	F1	F2
Caso A	0,878	0,001
Caso B	0,913	0,013
Caso C	0,887	0,080
Caso D	0,609	0,378
Caso E	0,300	0,677
Caso F	0,873	0,073
Caso G	0,961	0,021
Caso H	0,993	0,004
Caso I	0,359	0,403
Caso J	0,014	0,859
Caso K	0,568	0,401
Caso L	0,456	0,517
Caso M	0,054	0,585
Caso N	0,543	0,016
Caso O	0,908	0,075
Caso P	0,018	0,923
Caso Q	0,259	0,713
Caso R	0,739	0,223
Caso S	0,925	0,010
Caso T	0,964	0,006

3.3 Production program

It is contemplated the realization of a cycle of the cultivation of the production every 6 months of the Tilapia, considering more experience and greater capacity of infrastructure like of organisms of 50 org. / M3 to 100 org. / M3. Considering 20% mortality, with an initial weight of 5 g to reach a weight of 500 g average between 5-6 months where for its calculation we will use the following mathematical model.

Crop Gain = ((Number of Alevin * 80) / 100) * 40
Where the Gain of Farming is going to be equal to the amount of fry 80% of the quantity sown by concept of mustarded by the price of the market that is 40.

Table 8. Production.

Case	Quantity of alevines	Investment of Project	Gain of Culture
Caso A	1000	24160	32000
Caso B	1450	33885	46400
Caso C	1900	57610	60800
Caso D	2800	77060	89600
Caso E	6500	127410	208000
Caso F	1700	45710	54400
Caso G	2500	66210	80000
Caso H	3300	100710	105600
Caso I	4900	141710	156800
Caso J	11400	233460	364800
Caso K	3600	88860	115200
Caso L	5350	130935	171200
Caso M	7100	187010	227200
Caso N	10600	271160	339200
Caso O	24700	467610	790400
Caso P	7800	162160	249600
Caso Q	11650	240885	372800
Caso R	15500	333610	496000
Caso S	23200	491060	742400
Caso T	54100	907910	1731200

Each month and a half should be made with the same amount of fry as the applied case, but the investment process will be less because you will not have to invest in infrastructure only in fry to plant and food, giving a higher profit margin at the time of trading after the second cycle, assuming an invariable sale price of mxn 40.00 per kilogram.

4 Conclusions

Aquaculture projects are beginning to take place in many places, where the main obstacle to their success is the lack of knowledge of the basic principles and the necessary technical skills. In Mexico, the projection of an aquaculture farm requires models established in other countries, making the necessary modifications according to the natural conditions in the country. The development of aquaculture in Mexico compared to the world aquaculture development shows a lag in both the diversity and use of resources as in the modernization of the sector. Mexico is in an ideal situation for the cultivation of several species because it has extensive coastlines, abundant wetlands and optimal climates and fish farming has become a necessity in many places where this activity was not previously practiced.

Tilapia farming plays a crucial role in food security and nutrition as a nutrient-rich food source. It also provides significant jobs and income in the rural population, this work tries to demonstrate the feasibility of implementing a tilapia breeding in Mexico where natural conditions do not affect the crop, and serves as a guide to take decisions to The time to develop a project that has specific characteristics and has limited resources and can take full advantage of its development giving the investor the opportunity to know which is the most appropriate model depending on your budget and location.

References

1. J. L. Vicente, Análisis de componentes principales, vol. 32, Cataluña, 2002.
2. L. E. Pla, Análisis multivariado: método de componentes principales, Washington, D.C, 1986, pp. 87-89.
3. Aluja T, Morineau A. Aprender de los datos: el análisis de componentes principales. Una aproximación desde el Data Mining. Ediciones Universitarias de Barcelona. Barcelona. 1999.
4. T. A. Banet y A. Morineau, Aprender de los datos: el análisis de los componentes principales: una aproximación desde el" data mining, España: EUB, 1999.
5. Tierra Fertil, 2014. [On line]. Available: http://www.tierrafertil.com.mx/cuanto-cuesta-instalar-un-criadero-de-tilapias/. [Accessed: April 28, 2017]
6. Mercado libre mexico, 2017. [On line]. Available: http://listado.mercadolibre.com.mx/geomembrana-tangeomex. [Accessed: April 28, 2017]
7. Academia, 2014. [On line]. Available: http://www.academia.edu/9665033/SERVICIOS_Y_VALORES_INTEGRADOS_AL_PROYECTO http://www.academia.edu. [Accessed: April 28, 2017]
8. SAGARPA, 2009. [On line]. Available: http://www.gbcbiotech.com/genomicaypesca/documentos/peces/tilapia/Propuesta%20para%20desarrollo%20del%20cultivo%20de%20tilapia%20en%20Mexico.pdf. [Accessed: April 28, 2017]
9. F. Cantor, Manual de produccion de tilapia, Puebla, Mexico, 2008.
10. M. A. Saavedra , «Introducción al Cultivo de Tilapia,» Managua, Nicaragua, 2003.
11. M. B. Castañeda, A. F. Cabrera y Y. Navarro, Procesamiento de Datos y Analisis Estadisticos Utilizando SPSS, Porto Alegre, Brasil: EDIPUCRS , 2010.