Estimate of the cost of land along the ByPass in 2021 by the Gompertz diffusion model, using the Gauss-Newton algorithm.

RAMBININTSOA Franco Patrick, RAKOTO DAVID Séraphin, RAMBININTSOA Tahina, NDRIANAMBININA Djohary, HENIPANALA Mampionona

Université d'Antananarivo - Ecole Supérieure Polytechnique d'Antananarivo– Mention Urbanism Architecture and Civil Engineering – BP 1500 Antananarivo 101 Madagascar -francrambi@yahoo.fr

Abstract

The Malagasy State has adopted different land policies to better support its development of the territory. This publication estimates the cost of land along the By-Pass for 2021. To do this, we use Gompertz's non-linear mathematical model, which then uses the Gauss-Newton algorithm.
This simulation is intended for investors and the public sector to reduce land speculation in the city of Antananarivo.

Keywords : Nonlinear Mathematical Model, Land Security, Antananarivo, By-Pass

1- INTRODUCTION

Madagascar is among the countries in difficulty in the world. The Malagasy state has adopted various land policies with the main objective of meeting the massive demand for land security, which is the basis for the development of the territory.
For Antananarivo, urbanization is spread along the major roads. In this scientific publication, we will estimate the cost of land along the By-Pass for 2021. This modelling will be used as a calibration of the forecast of the cost of land along the tracks of the new roads built in Antananarivo, whose urbanization plans are being developed.

2- LITERATURE REVIEW

- Bjorck, A. (1996). Numerical Methods for Least Squares Problems. SIAM. Philadelphia, PA.
- Dulmage, A.L. and N.S. Mendelsohn (1963). Two algorithms for bipartite graphs. SIAM 7, 183–194.
- Gilli, M. (1995). Graph-Theory Based Tools in the Practice of Macroeconometric Modeling. In : Methods and Applications of Economic Dynamics (S. K. Kuipers, L. Schoonbeek and E. Sterken, Ed.). Contributions to Economic Analysis. North Holland. Amsterdam.

3- METHODS - GENERALITY ON NON-LINEAR MODELS

3.1-Definition of Non-Linear Regression Models and Estimation Methods

A model can be non-linear in both variables and parameters. If its parameters are non-linear and the explanatory variables are linear, the model is said to be intrinsically non-linear (difficult to linearize, we use certain algorithms, in particular that of Gauss-Newton). On the other hand, an intrinsically linear model (i.e., a non-linear model that can be linearized after certain transformations: logarithmic, etc.) is one where all parameters are linear, even if the explanatory variables are non-linear.
In terms of estimation methods, the following are noted:

- For intrinsically linear models, OLS can be applied to the linear forms obtained after transformation. These are the "easy linearization non-linear models".

- For complex non-linear models, they are quite difficult to linearize.
- For inherently non-linear models, we use general non-linear regression methods that require iterative numerical procedures, in particular: the Gauss-Newton algorithm, that of Newton-Raphson, the Marquard algorithm, ...).

Table1 - Non-linear models with easy linearization

Non-linear models	Caracteristics
Log-log or double log models	presence of logarithm in both members of the equation (model)
Semi-log models	Presence of logarithm in a single member of the equation so that one member remains linear and another taken in logarithm: either to the left (log-lin model) or to the right (lin-log model).
Reciprocal models	Presence of terms (variables) in inverse or reciprocal form in the model. Example: inverse relationship inflation and unemployment.
Log-inverse or log-hyperbole or log-reciprocal models	These are both reciprocal and semi-log models where the dependent variable is taken in logarithm and one or more explanatory variables are taken in inverse form
Polynomial models	Presence of a single explanatory variable in the right limb with different powers.

These non-linear models with linearization possibilities [2]

3.2- complex non-linear models

Complex non-linear models are models for which linear transformation is difficult. We distinguish:
- The dissemination models translating the “market life curve” (land) models are an illustration of complex non-linear models. These concepts use fairly complex mathematical formulations.
- Dissemination models that explain the “stacking of market sales” (of land) over time.
Graphically (Figure 1), these two models are as follows:

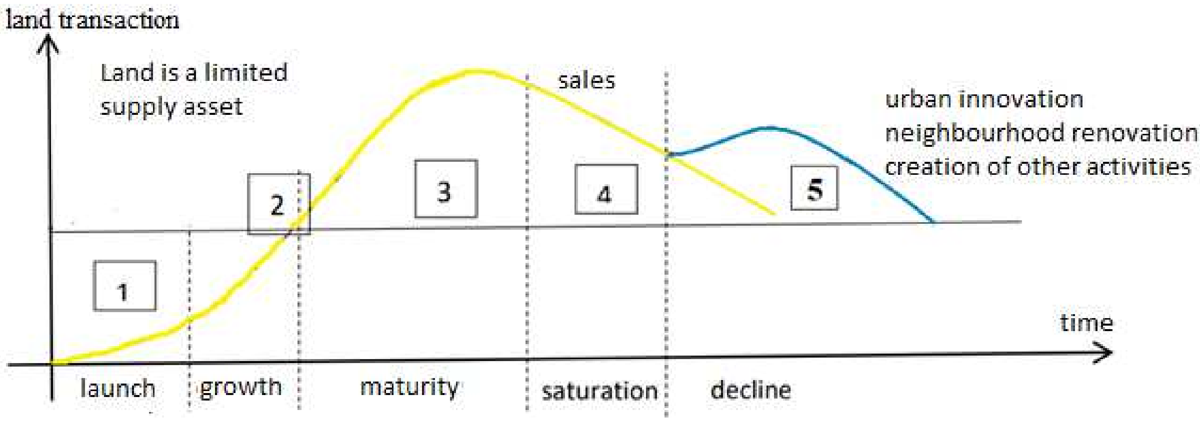

Figure 1 – Land Change Curve

Interpretation of the Real Estate sales curve:

This curve represents the Land Market in a newly urbanized area such as the By-Pass, as a function of time. It reflects the reality of transactions and it is difficult to find a mathematical expression that translates it. To better understand the explanation of the curve, we take the case of the By-Pass, which was built in 2005.
Note that land is considered an "Asset" but not an "asset". Its supply is limited by space, and once consumed, we can only produce it through demolition or "rebuilding the city on its own".

Phase 1 «launch»: this is the number of land sales around the beginning of the work This phase includes speculations in all forms. For land sales near the By-Pass, this phase was extended from 2004 to 2010. Most of the land along the By-Pass was still undestructible during this period. Only a northern part of the new track was buildable.

Phase 2 "Growth" is the number of transactions between 2011 and 2017 for the By-Pass. An Operationalization of the Antananarivo metropolitan area PUDi was carried out from 2011 to 2014 and the By-Pass areas were affected. Two (02) PUDEs were developed in 2012 along the By-Pass. However, the by-laws of these planning documents were not enforced. Derogations from "backfill permits" and "building permits" were deliberate within the "Ad'Hoc Committee" of the Ministry of Spatial Planning from 2011 to 2014. This increased transactions along the By-Pass. In addition, the illegal embankment proliferated from 2012 to 2017. These reasons explain this second phase of the curve.

Phase 3 "Maturity": this is the current phase of land transactions along the By-Pass. In 2017, the By-Pass was approved and is being applied. The illegal embankment becomes a criminal case in the court. This PUDé freed up 420ha of land along the BY-Pass for urbanization. This increased the number of land sales along the By-Pass.
We can see the land owned by large landowners (area > 1 hectare) backfilled along the track. These urban developments explain the current increase in the price of land in the area. In 2021, we are in the vicinity of the peak of the transaction curve. The following facts justify reaching this peak:

- Current land owners are more or less stable (i.e., they do not speculate). These are large landowners, businesses, wealthy families, etc.).
- Most of the land along the track is already sold to these stable owners. And since the supply of land is very limited, there will be no new land on the land market.

These facts explain that the number of land transactions along the By-Pass would hardly increase. Hence the peak of this third phase in 2021.

NB: there is a decrease in the number of sales (or transactions), but not in the price of land. The cost of land is a growing curve with the appearance of an exponential function.

Phase 4 "Saturation" and Phase 5 "Decline" do not concern the number of land transactions at the By-Pass zone level. This part does not yet concern the Forecast, so we do not predict the years of the “saturation and decline” phases for the By-Pass.

We can see the case of the residential area of Ivato for the phases of saturation and decline in the number of land transactions. There is an increase in the price of land in these phases the demand is much higher than the supply which is very limited.

Innovation phase: this phase is to be conceived by urban planners and economists after or during the phase of decline in the number of land transactions. This phase of innovation is the consequence on land of the restructuring of precarious neighborhoods and the creation of urban attractions. This phase creates a revitalization of space through new zoning policies and a reconstruction of the city on its own that is to say an urban densification within the neighbourhood. [4]

We have already explained that non-linear models are difficult to estimate mathematically. Among the complex models, here are some often used for solving real problems. Before detailing these models, we recall that they are used in several fields of research such as hydrology (ex: Caquot formula in urban sanitation), econometrics (ex: inflation relationship – unemployment) and biology (ex: evolution of bacteria)...

3.3- The non-linear Gompertz model applied to the valuation of land costs

Its equation is as follows:: $\frac{dN}{dt} = - rN\, lnN$ with r > 0

Using the Gompertz model, the change in the cost of land follows the following mathematical law:

$\frac{dC}{dD} = rCln\frac{Z}{C}$

With : C : cost of land

D : distance from the development pole

Z : parametric value associated with zoning

r : rate of growth of the price of land.

C_0 : cost of land at development pole

$$\frac{dC}{dD} = rCln\frac{Z}{C} \qquad (**)$$

$$\frac{dC}{Cln\frac{Z}{C}} = r\, dD$$

$$\frac{dC}{C(lnC-lnZ)} = -\, r\, dD$$

By posing u = lnC – lnZ , we have $\frac{du}{dC} = \frac{1}{C}$

By calculating the primitives of the two members, we obtain:

$$\int \frac{1}{C(lnC-lnZ)} dC = \int -r\, dD$$

⇨ $ln(lnC\text{-}lnZ) = \text{-}rD + constant$

$lnC – lnZ = \lambda e^{-rD}$

$ln\, C = lnZ + \lambda e^{-rD}$

$$C = e^{\ln Z + \lambda e^{-rD}}$$

$$C = Z\, e^{\lambda e^{-rD}} \qquad [3]$$

Let us now calculate the constant λ by the initial condition $C(0) = C_0$

$C_0 = Ze^{\lambda} \Rightarrow \lambda = \ln \frac{C_0}{Z}$

As a result of the equation (**), we obtain: $\mathbf{C = Z.\, e^{\ln\frac{C_0}{Z}.e^{-rD}}}$

4- FINDINGS : CALCULATION OF C_0 AND RATE R OF THE GOMPERTZ MODEL BY THE GAUSS NEWTON ALGORITHM

4.1- Reminder of the Gauss-Newton algorithm

This resolution uses very complex iterative numerical procedures. The idea is to arrive at new coefficients that allow to obtain a linear form, by proceeding to a limited development of Taylor on the basis of initial coefficients or iterative values. Then, apply the OLS to the linear form thus obtained to estimate the new coefficients. Iterations stop more and more as the coefficients become stable. However, "estimator convergence" is a criterion that sanctions the efficiency of the method used (initial values should be closer to optimal values).

Let the non-linear model: $Y_t = (X, a) + e_t$ (1)

With : Y_t = dependent variable (in matrix form)

X_t = vector of explanatory variables (matrix of observations of explanatory variables of dimension « n, k + 1 »)

a = parameter vector to be estimated (dimension « k + 1 »)

e_t = random term or error.

The least squares estimator for the non-linear model (1) is "a" which minimizes the « $(a) = e'e = [Y_t - (X, a)]' [Y_t - f(X, a)]$ » as follows (we have k+1 conditions of the 1st order: $\partial S/\partial a = 0$) :

$\frac{\partial S}{\partial a} = -2 \frac{\partial f(X,a)}{\partial a} [Y_t - f(X, a)] = 0$

With : $\frac{\partial f(X,a)}{\partial a} = Z(a) = \begin{pmatrix} \frac{\partial f(X_1,a)}{\partial a_0} & \cdots & \frac{\partial f(X_1,a)}{\partial a_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(X_n,a)}{\partial a_0} & \cdots & \frac{\partial f(X_n,a)}{\partial a_k} \end{pmatrix}$ the jacobian matrix J(x).

Note: $(a*)$ = matrix calculated for the particular values of a ($a = a*$). Due to Taylor's limited development in the vicinity of « $a*$», we obtain the following approximation to the «t-th observation»:

$f(x_t, a) \simeq f(x_t, a) + \{\frac{\partial f(X_t,a)}{\partial a_0}|_{a = a^*} \dots \dots \frac{\partial f(X_t,a)}{\partial a_k}\}(a - a^*)$

$\{\frac{\partial f(X_t,a)}{\partial a_0}|_{a = a^*} \dots \frac{\partial f(X_t,a)}{\partial a_k}\}$ is the matrix of the first derivatives noted ∇f(x) or ∇r(x)'r(x)

Otherwise, in matrix notation:

$f(X, a) \simeq$ f(X,a*) + Z(a*)(a-a*)… … (2)

(2) may also be written:y = f(X,a*) + Z(a*)(a-a*) + e

Ou y = f(X,a*) + Z(a*)a - Z(a*)a* + e … … (3)

By posing: $\bar{y}$(a*) = y - f(X,a*) + Z(a*)a* , then (3) is reduced to:

$\bar{y}$(a*) = Z(a*)a + e …. …. (4)

The estimator of this “linear model (4) “ (by OLS) is:

$a^2 = [Z(a^*)' Z(a^*)]^{-1} Z(a^*)' \bar{y}(a^*) = a^* + [Z(a^*)' Z(a^*)]^{-1} Z(a^*)'[y-f(X,a)]$

We will have « k+1 » new values for the vector « $a = a$ Convergence is achieved when « $\hat{a} = a^p \simeq a^{p-1}$ » (stability of coefficients in the p-th iteration).

4.2- Estimate of the cost of land along the ByPass for the year 2021

Figure 2 shows the Zoning of the PUDé By-Pass 2018

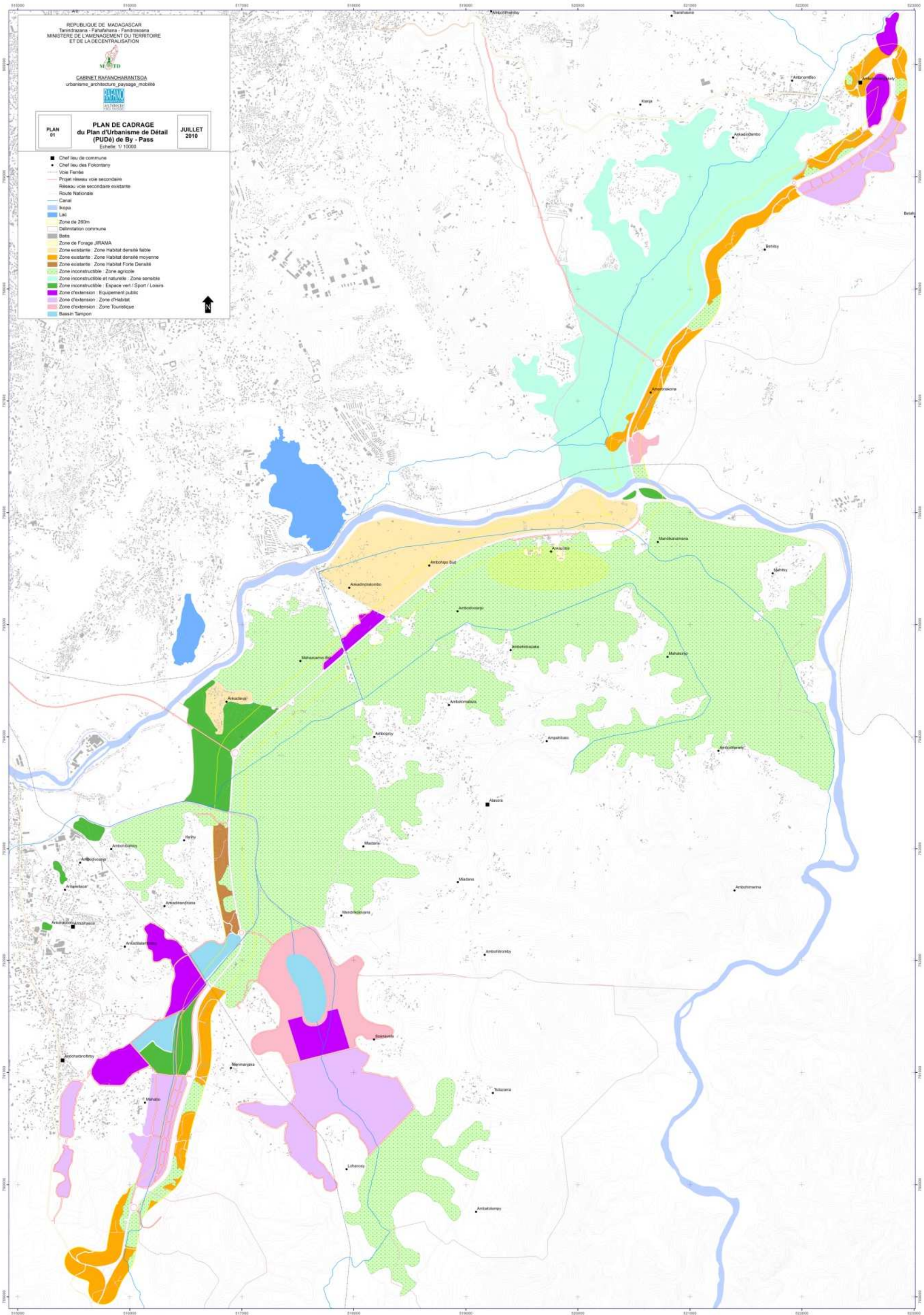

Figure 2 – Zoning of the PUDé By-Pass 2018

Remember that we have the equation: $\mathbf{C} = \boldsymbol{Z}.\boldsymbol{e}^{\ln\frac{C_0}{Z}.e^{-rD}}$

Table 1 : Cost of land along the ByPass for the year 2021

D in kilomètre	Z	C in thousand of Ariary
0,25	2	360
0,25	3	377
0,25	4	390
0,75	2	322
0,75	4	352
1	2	310
1	4	337
1	3	325
1,5	1	257
1,5	3	300
2	1	235
2	3	277
2	4	290
2,5	2	243
2,5	3	260
3	1	200
3	3	243
3	4	255
3,5	2	213
3,75	4	235
4	1	175
4	2	200
4,5	3	205
4,75	4	210
5	1	155
5	2	180
5	4	205
6	2	160

Here is the Matlab script for calculating the coefficients:
NB : to avoid any ambiguity, the variable "d" is replaced by "q" in this Matlab script.

```
clear all ; close all ; clc ;
%%%%%%%%%%%%%%%%%% Gauss Newton Gompertz%%%%%%%%%%%%%%%%%
idelta(1,1) = 450 ; idelta(2,1) = -0,1 ;

qi = [0.25  0.25   0.25   0.75   0.75    1     1    1     1.5    1.5     2     2     2     2.5    2.5    3...
      3     3     3.5    3.75    4      4    4.5   4.75   5      5      5     6];
```

```
zi = [  2     3      4      2     4      2     4     3      1      3      1     3     4      2     3     1...
        3     4      2      4     1      2     3     4      1      2      4     2];
ci = [360   377   390   322   352   310   337  325  257   300   235  277  290   243   260  200...
       243   255   213   235   175   200   205  210  155   180   205  160];

err = 0.3. *randn(1,numel(yi)) ; it = 0 ;
%%%%%%%%%%% FORMAL DERIVATION CALCULATION %%%%%%%%%%%%%%
syms p1 p2 q z
modelGompertz = z. * exp(ln(p1/z).exp(-p2.q)) ;

dy_d1 = diff(modelGompertz,p1) ; dy_d2 = diff(modelGompertz,p2) ;
dy_d11 = subs(dy_d1,q,qi,z,zi) ;  dy_d22 = subs(dy_d2, q,qi,z,zi) ;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 itmax = 100 ; tol = 1e-10 ;
while it < itmax

it = it + 1 ;
 dy_da= subs(dy_d11, { 'p1 ' , 'p2 ' } , {idelta(1)   idelta(2)});
 dy_db= subs(dy_d22, { 'p1 ' , 'p2 ' } , {idelta(1)   idelta(2)});

    Jacob(1:numel(xi),  1) =  dy_da ;  Jacob(1:numel (xi),  2) =  dy_db ;
 F = yi - subs(modelGompertz, { ' q ' , ' z ' ,'p1 ' , 'p2 ' } , ...
      {qi   zi    idelta(1)    idelta(2) });

ndelta = inv(Jacob' *Jacob) *(Jacob' *F')  +  idelta ;  idelta = ndelta ;
fx = subs(modelGompertz,{ ' q ' , ' z ' ,'p1 ' , 'p2 '}, ...
      { qi   zi    idelta(1)   idelta(2)})  ;

Fn(it)  =  sum(F) ;
rmse(it+1)  =  sqrt(Fn(it).^2)/(length(ci) - length(idelta))  ;

If  rmse(it + 1) - rmse(it)  <  tol
         sol  =  idelta ;
    break
end
```

Thus the result obtained by simulation is: $C = Z.e^{\ln\frac{396}{Z}.e^{-0,0347D}}$

5- DISCUSSION

We note that the linear system that defines the Gauss-Newton step is a system of normal equations and that the step is also the solution of a least squares problem. The algorithm may not converge if the starting point is chosen too far from the solution. When the residues are large at the point of the solution, the approximation of the second derivative matrix may be insufficient, resulting in either very slow convergence or no convergence at all.

6- CONCLUSION

Au terme de cette publication, il est nécessaire de rappeler que la sécurisation foncière joue un rôle plus que prépondérant sur le développement économique pour le cas de Madagascar. Le foncier urbain offre des opportunités de création d'emploi et de promotion du logement, en particulier pour Antananarivo. Cette recherche nous a donné la modélisation du coût du foncier longeant le By-Pass pour l'année 2021. Cette simulation intéresse les investisseurs et le Secteur publique et évite toute spéculation foncière. Notre résultat peut servir de calage pour d'autres évaluations foncières pour les nouvelles voies et pour les autres villes.

BIBLIOGRAPHICAL REFERENCES

- [1] J. E. Kelley, Jr. The cutting-plane method for solving convex programs. Journal of the Society for Industrial and Applied Mathematics, 8 :703–712, 1960.
- [2] R. T. Rockafellar. Convex analysis. Princeton university press, 2015.
- [3] C. Lemaréchal and F. Oustry. Nonsmooth Algorithms to Solve Semidefinite Programs. In L. El Ghaoui and S.-I. Niculescu, editors, Advances in Linear Matrix Inequality Methods in Control, chapter 3. SIAM, 2000.
- [4] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM journal on imaging sciences, 2(1) :183–202, 2009.
- [5] B. Hiriart-Urruty and C. Lemaréchal. Convex analysis and minimization algorithms. II, volume 306 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, 1993. Advanced theory and bundle methods

ACKNOWLEDGMENTS

In appreciation, this publication's authors want first to give thanks to the Lord, without whom nothing of the smallest details of our lives would have taken place. His grace has provided us with faith, strength, health needed for this research, and to reach the terms of this publication, which is the result of several years of research in the Polytechnic School Antananarivo that we express our deep gratitude and especially its director, Professor Rijalalaina Rakotoson.

We are also thankful to the Editorial Board of the Madarevues And Mada-Hary, as well as their team. Special thanks then go to the Alasora Community for their cooperation during the development of this scientific publication.

We also need to thank all those who have directly or indirectly contributed to this publication's realization. Last but not the least we do not forget especially, our family who was always at our side during the development of our research.

The authors.