A Model of Student Flow through the College Curriculum

Oscar Duarte, Ph.D.1 and Christian Márquez, M.Sc[1]
[1] Universidad Nacional de Colombia, Colombia, ogduartev@unal.edu.co,cdmarquezb@unal.edu.co

Abstract*—This paper presents a mathematical model of the flow of students enrolled in a college program through the curriculum. From a simplified model of the curriculum it is possible to estimate the number of students in every level, number of graduates and graduation times of a specific program or a set of programs. Some related phenomena, such as cancellation of courses, repetition, approval, registrations, transfers (and dual degrees programs), loss of student's quality and graduation are modeled. The model has been validated with historical data from the School of Engineering of the National University of Colombia. The model analysis and several simulations allow to understand the effects of the modeled phenomena on the total number of students in a curriculum and on their graduation times. Using the model there is also a partial explanation regarding the increase in the number of students registered at the School of Engineering since 2009.

Key words - Student flow, Curriculum model

Digital Object Identifier
(DOI):http://dx.doi.org/10.18687/LACCEI2016.1.1.214
ISBN: 978-0-9822896-9-3
ISSN: 2414-6390

A Model of Student Flow through the College Curriculum

Oscar Duarte, Ph.D.[1] and Christian Márquez, M.Sc[1]

[1] *Universidad Nacional de Colombia, Colombia, ogduartev@unal.edu.co,cdmarquezb@unal.edu.co*

***Abstract*—This paper presents a mathematical model of the flow of students enrolled in a college program through the curriculum. From a simplified model of the curriculum it is possible to estimate the number of students in every level, number of graduates and graduation times of a specific program or a set of programs. Some related phenomena, such as cancellation of courses, repetition, approval, registrations, transfers (and dual degrees programs), loss of student's quality and graduation are modeled. The model has been validated with historical data from the School of Engineering of the National University of Colombia. The model analysis and several simulations allow to understand the effects of the modeled phenomena on the total number of students in a curriculum and on their graduation times. Using the model there is also a partial explanation regarding the increase in the number of students registered at the School of Engineering since 2009.**

***Index Terms*—Student flow, Curriculum model**

I. Introduction

Every Higher Education institution faces the problem of predict the number of students in every level. The prediction is crutial, for instance, to decide the number of courses to be offered and the number of professors required. For steady-state conditions, the expected number of students is constant, therefore the number of courses and professors are also constant. Some mathematical models use this static approach to estimate the resources requirements (e.g. [1]).

The student flow is the way an student (or a group of student) pass from one education level to another every academic term. There are descriptive studies about the historical student flow of particular institutions from an statistical point of view (e.g. [2]). The number of students in a specific level at a specific academic term depends on the number of students in previous levels and terms. From this point of view, student flow is a dynamic system that can be modeled using difference equations. This approach is used in [3].

A dynamic model is very useful to analyze and predict short and medium term phenomena. As an example, consider the case of the School of Engineering of the National University of Colombia. It has had a continued growth in its offer of academic services. Figure 1 1 shows the evolution of the number of programs offered by the School since 1980. Related to the growth of academic offerings is the growing number of active students. Figure 2 shows the evolution of the number of students of the School of Engineering from 2009 to 2013. Figure 2a shows a growth of approximately 1.000 students in 4 years. Figure 2b shows that this growth is mainly due to the increase in the number of undergraduate students.

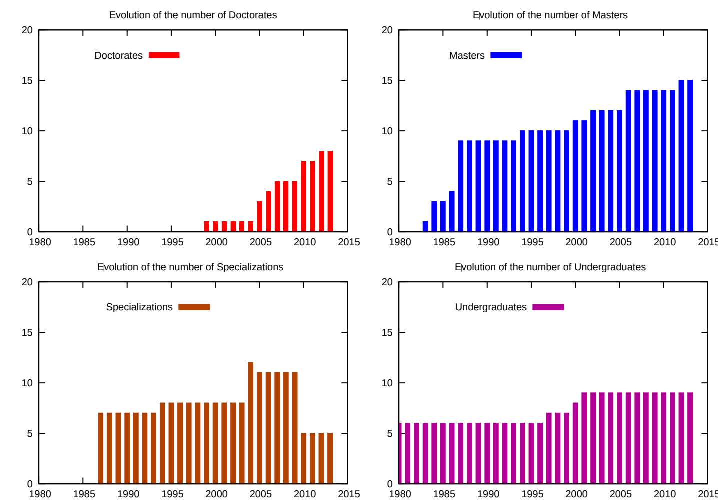
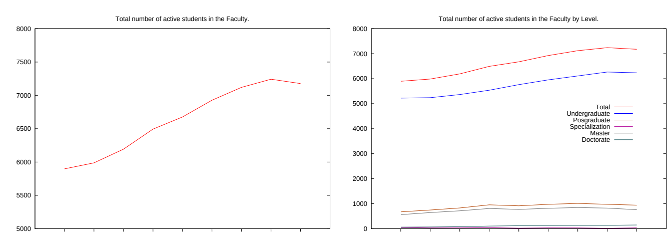


Figure 1. Evolution of the number of programs offered by the School of Engineering at the Universidad Nacional de Colombia.



(a) Total (b) By level

Figure 2. Evolution of the number of students in the School of Engineering at the Universidad Nacional de Colombia.

This growth results with it pressure on the demand for courses, workspaces, professors, among others, that the institution must cover. For this reason, it is necessary to have a model to help predict the future demand to be covered. Moreover, in 2008 an Academic Reform was applied ([4]), and it is important to assess its effects. One of the many perspectives from which it should be evaluated is to determine whether or not it has had an impact on the number of students and their permanence time in college.

This paper presents a mathematical model of the flow of students enrolled in college program through the curriculum. Our first aim was to understand the causes of the behavior shown in 2, but the model may be used in a broader scope.

Digital Object Identifier (DOI): http://dx.doi.org/10.18687/LACCEI2016.1.1.214
ISBN: 978-0-9822896-9-3
ISSN: 2414-6390

The model allows to simulate the impact of various academic aspects that changed in the 2008 reform. The model is built using tools for modeling, analysis and simulation of discrete dynamical systems. The approach is similar to the model of [3]; however, a more detailed mathematical analysis is conducted and more parameters and phenomena are analyzed. This document presents the following results:

- A model that quantifies the effect on the *number of students* and the *time of graduation* of:
  - Approval rate.
  - Cancellation rate[1].
  - Dessertion rate.
  - Registrations.
  - Transfers (and double degree programs).
- The parameters associated with the model adjusted to the case of the School of Engineering.
- A partial explanation ($\approx$ 85%) of the causes of the increase in the number of undergraduate students in the School.

There is a lot of work about the factors that impact on graduation, retention, graduation times and other related phenomena (e.g. [5]–[9]). Notice that this paper is not focused on that important aspect but in the mathematical relationships of those phenomena.

The paper is organized as follows: Section IIshows the mathematical model and a reduced model, Section III presents the model validation exercise with actual historical data associated with the School of Engineering, in Section IV shows several exercises in mathematical analysis and numerical simulation of the reduced model approach to understanding the dynamics of the modeled phenomena, in Section V the model is used to find the causes of growth in the number of undergraduate students who are shown in figure 2; the conclusions are given in section VI.

II. The mathematical model

The undergraduate curriculum at the School of Engineering consists of approximately 60 courses between mandatory, optional and free choice. The curricula are organized through curriculum nets that distribute these courses into 10 semesters. The mathematical model is a considerable simplification of these curricula.

Figure 3 shows the curriculum of the proposed model. It is a curriculum of n academic semesters with a single course in each semester which is a prerequisite for the course in the following semester. This model can be interpreted as the ‘backbone’ of a real curriculum.

On the proposed plan of study the following dynamic behavior is formulated and illustrated in Figure 4:

D.1. Subscript j is used to distinguish each of the n courses in the curriculum. Because in each academic semester there is a single course, subscript j also serves to identify the semester.

[1]A ‘cancellation’ occurs when, at first, a student is registerd in a course but then he/she cancells that registration.

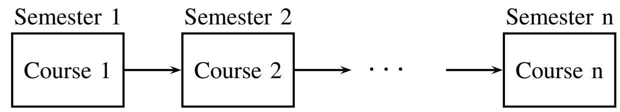


Figure 3. Curriculum of the minimal model

D.2. Variable k is used to refer to the so-called ‘academic term’ or just ‘term’. There are 2 ters in every year[2]. k is a discrete variable that can be assimilated to the sequence of terms or integers ($k = 1, 2, \cdots$).

D.3. The number of students enrolled in the course j for the academic term k is denoted $x_j(k)$.

D.4. When passing from one academic term to the next one, some students go to the course of the next semester, some students do not succeed, either by canceling or because they do not approve (see Figure 4):

$$x_{j+1}(k+1) = x_{ap_j}(k)+x_{can_{j+1}}(k)+x_{nap_{j+1}}(k)+t_j(k) \tag{1}$$

Where:

- $x_{ap_j}(k)$: Number of students that approved course j for term k.
- $x_{can_{j+1}}(k)$: Number of students that cancelled course $j+1$ for term k.
- $x_{nap_{j+1}}(k)$: Number of students that registered and did not cancelled nor approved course $j+1$ for term k and that neither abandon the program.
- $t_j(k)$: Number of students who entered the program due to transfer o double degree authorization. These students begin clases in term $(k+1)$ registering course j.

D.5. For the $1st$ term the dynamic is represented by:

$$x_1(k+1) = x_{can_1}(k) + x_{nap_1}(k) + u(k) + t_j(k) \tag{2}$$

Where:

- $u(k)$: Number of students admitted for term k who begin their studies in term $(k+1)$

D.6. The following parameters are defined:

- $\alpha_j(k)$: Approval rate for course j for term k.
- $\beta_j(k)$: Cancellation rate for course j for term k.
- $\sigma_j(k)$: Desertion rate for j for term k.

D.7. The right side of (1) can then be calculated this way:

$$\begin{aligned}x_{ap_j}(k) &= x_j(k)(1-\beta_j(k))\alpha_j(k)(1-\sigma_{j+1}(k))\\ x_{can_{j+1}}(k) &= x_{j+1}(k)\beta_{j+1}(k)(1-\sigma_{j+1}(k))\\ x_{nap_{j+1}}(k) &= x_{j+1}(k)(1-\beta_{j+1}(k))(1-\alpha_{j+1}(k))(1-\sigma_{j+1}(k))\end{aligned} \tag{3}$$

D.8. The number of students who enrolled in a semester equal or minor to j during term k is denoted by $v_j(k)$.

D.9. The total number of students in the program during term k, $y_T(k) = y_n(k)$, is denoted by $y_T(k)$.

[2]We use to name them ‘2009-01’, ‘2009-02’, ‘2010-01’, etc.

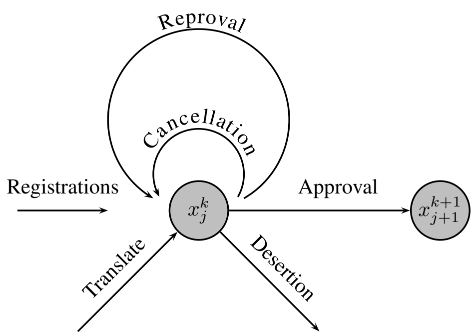


Figure 4. Flow diagram of the state $x_j(k)$

D.10. The total number of students who graduate by the end of academic term k is denoted by $g(k)$.

With the conditions below it can be derived a dynamic discreet lineal model that varies in time:

$$\begin{aligned} X(k+1) &= A(k)X(k) + B(k)U(k) \\ Y(k) &= C(k)X(k) + D(k)U(k) \end{aligned} \tag{4}$$

A. Reduced model

With the purpose of facilitating the analysis of the model, two simplifications are introduced:

- Non variability in time: All the parameters of the model remain constant through time.
- Homogeneity: The parameters of the model are equal for n courses of the curriculum.

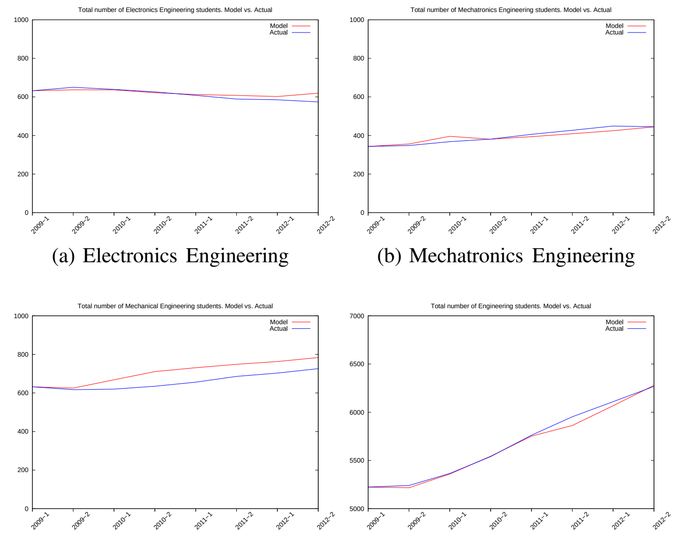
The former simplifications imply:

$$\begin{aligned} \alpha_j(k) &= \alpha \quad \forall j,k \\ \beta_j(k) &= \beta \quad \forall j,k \\ \sigma_j(k) &= \sigma \quad \forall j,k \end{aligned} \tag{5}$$

The reduced model results in a lineal and non-variable model through time (all proofs are in [10]):

$$\begin{aligned} X(k+1) &= AX(k) + Bu(k) \\ y(k) &= CX(k) + Du(k) \\ A &= \begin{bmatrix} \delta & 0 & \cdots & 0 & 0 \\ \gamma & \delta & \cdots & 0 & 0 \\ 0 & \gamma & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \vdots & \gamma & \delta \end{bmatrix}_{n\times n} \quad & B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1} \\ C &= \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}_{1\times n} \quad & D = \begin{bmatrix} 0 \end{bmatrix}_{1\times 1} \\ \gamma &= (1-\beta)\alpha \\ \delta &= 1-\gamma \end{aligned} \tag{6}$$

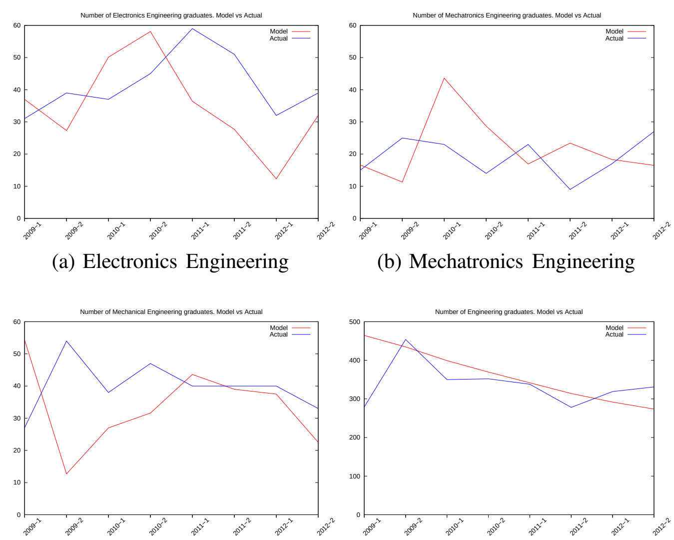
For a 10 semester's program, considering 8 terms, this simplification reduces the number of parameters from $10 * 8 * 3 = 240$ to only 3. In this document it will be used the complete model in the validation and use exercises which involve data from the School of Engineering in sections III and V.



(a) Electronics Engineering (b) Mechatronics Engineering

(c) Mechanical Engineering (d) School

Figure 5. Validation using the total number of students



(a) Electronics Engineering (b) Mechatronics Engineering

(c) Mechanical Engineering (d) School

Figure 6. Validation using the number of graduates

The reduced model will be used in the analysis exercises in section IV. The two models can be furthermore separated from each other due to the use or not of the sub-indexes and the variability through time $j(k)$ in the parameters.

III. Model Validation

To validate the model two procedures were carried out. The first one with information from three programs (Electronics, Machatronics and Mechanical Engineering) and the second one with information of the School as a whole. The results are shown in Figure 5.

The procedure for the 3 programs was the following:

Vc.1. For each program the courses of the active curriculum that could represent better the simplified model of

Figure 3 were selected. In this task there was companionship from the curricular coordinators of the selected programs.

Vc.2. Databases were designed (using a `mysql` storage engine) to organize the primary information given by the Admissions (Registrarion) Office (See section III-A).

Vc.3. The primary information was processed in order to estimate the parameters, the initial conditions and the entry variables of the model, as detailed in section III-B.

Vc.4. The mathematical model was implemented using a simulation tool (`scilab`).

Vc.5. The model was ran with the parameters and variables obtained in the previous step Vc.3. For estimating the evolution of the total number of students enrolled in each program from the 2009–1 term to the 2013–1 term.

Vc.6. There was a comparison between the total number of students estimated by the model and the real evolution experimented from the same term. The comparison is presented graphically in figures 5a, 5c, and 5c.

Vc.7. There was a comparison between the evolution of the total number of students who graduated estimated by the model and the actual evolution experienced during the same term. The comparison is shown graphically in figures 6a, 6c and 6c.

For the other hand, the validation procedure for the School as a whole was the following:

Vf.1. The primary information was processed in order to estimate the rate of loss of student's quality, the initial conditions and the entry variables of the model, as detailed in section III-B.

Vf.2. Due to the fact that the rates of cancellation and approval, as well as the initial conditions, are associated to the specific courses, and taking into account that for the validation process for the School, specific courses have not been chosen, these parameters were estimated in a different way. The procedure is explained in sections III-B8, III-B9 and III-B10.

Vf.3. The model was ran with the parameters and variables obtained in the Vf.1. and Vf.2 steps in order to estimate the evolution of the total number of students of the School from the 2009–1 term to the 2013–1 term.

Vf.4. There was a comparison between the evolution of the total number of students estimated by the model and the actual evolution registered during the same term. The comparison is shown graphically in figure 5d.

Vf.5. A comparison was made between the evolution of the total number of students who graduated estimated by the model and the real evolution registered in the same term. The comparison is presented graphically in figure 6d.

As a result of the validation processes it can be concluded that the model is able to properly reproduce the behavior of the total number of students, especially for the respective data of the School as a whole. The modeling of the number of students who graduated is less accurate, although the tendencies are indeed reflected in the model; again, predicting the number of students who graduated of the School as a whole is better than the prediction for each of the programs separated.

One of the reasons why is more difficult to model the number of students who graduated than the total number of students lies in the difficulty of estimating properly the initial conditions; in other words, how to distribute the total number of students in the initial academic term for the n semesters ($X(0)$).

A. Primary Information

For the estimation of the parameters and the entry variables the following primary information was used:

I.1. **Report of the active and blocked (non-active) students**. It is a chart generated by the Registration Office with information about every student of the School of Engineering. In it, it is reported the admission date to the academic program, the changes in the status (blockages in the academic record) and their causes. The chart includes the status changes since the 2006–1 term to 2013-1 term. The number of entries in the chart is 16297.

I.2. **Grading report**. It is a collection of charts generated by the Registration Office. In them are recorded the final grades obtained by every student of the School of Engineering in all the courses taken since the 2004-1 term to the 2013-1 term. The number of entries is 664115.

I.3. **Cancellation's Report**. It is a collection of charts generated by the Registration Office in which are reported the cancellations of courses previously registered. There are records since the 2004–2 term until the 2013–1 term. The total number of entries is 61781.

I.4. **Admissions Report**. It is a collection of charts generated by the National Direction of Admissions which contains the number of students admitted in each program. There are records since the 2007–1 term until the 2013-1 term.

B. Obtainment of parameters and entry variables

1) Approval Rates: Using the grades report it was calculated the approval rate for each of the courses of chart 1 from the 2009–1 term to the 2013–1 term. In each case there were considered only the records for students enrolled in the program of study. The approval rate of course j in term k was estimated like this:

$$\alpha_j(k) = \left.\frac{\text{Number of students with an approval grade}}{\text{Number of students with a reported grade}}\right|_{\substack{\text{course}=j \\ \text{term}=k}}$$

2) Cancellation Rates: Using the cancellation's report it was calculated the cancellation rate for each of the courses in chart 1 since the 2009-1 term to the 2013-1 term. In each case there were only considered the records of students enrolled in the program of study. The cancellation rate of course j in term k was estimated like this:

$$\beta_j(k) = \frac{\text{Number of students who cancelled}}{\text{Number of students with a reported grade}}\Bigg|_{\substack{\text{course}=j \\ \text{term}=k}}$$

3) Desertion Rates: Using the report of active and blocked students it was calculated $n_j(k)$ the number of active students in each program with $1, 2, 3 \cdots, n$ semesters of seniority. This calculation was made from the 2008-1 term to the 2013-1 term. The desertion rate of course j in term k was estimated like this:

$$p_j(k) = \frac{n_j(k) - n_{j+1}(k+1)}{n_j(k)}$$

The number of available data for each value of j is different. Consider, for instance, the case in which $j = 10$. The information of the number of students with 10 semesters of seniority can only be obtained in term 2013-1 and corresponds to those students that entered in the 2008–1 term. While for $j = 9$ there are two data: in the 2013-1 term are the students that entered in 2008-2 and in 2012-3 are the students who entered in 2008-1.

Moreover, in this estimation it should be included that cases of desertion for reasons different to the successful completion of the curriculum. As a result, the number of values for $p_j(k)$ that can be calculated is not the same for every value of j. If we denote this value by m_j the result is:

$$m_j = 11 - j$$

Therefore, it was decided to estimate the desertion rate as the medium value of the available data.

$$\sigma_j(k) = \frac{\sum^k p_j(k)}{m_j}$$

4) Registrations: Using the report of active and blocked students it was calculated $n_{1_{ex}}(k+1)$ the number of students active in each program with 1 semester of seniority in term $k+1$, who entered by admission exam. This number is interpreted as the number of students admitted in term k to start their studies in term $(k+1)$, which means

$$u(k) = n_{1_{ex}}(k+1)$$

5) Transfers: Using the report of active and blocked students it was calculated $n_{1_{tras}}(k+1)$ the number of active students in each one of the programs with a semester of seniority during term $k+1$, who entered by a different condition than an admission test (transfer, double degree program). This number is interpreted as the number of transfers in term k to start studies in the term $(k+1)$, in semester number $n/2 = 5$.

$$t_j(k) = \begin{cases} n_{1_{tras}}(k+1) & \text{si } j = 5 \\ 0 & \text{si } j \neq 5 \end{cases}$$

6) Initial Conditions: Using the grades report it was calculated $\tilde{x}_j(0)$ the number of students who took every one of the courses of chart 1 in the term 2009-1 ($k = 0$). Furthermore, using the report of active and blocked students it was calculated $N(0)$ the number of students active in 2009-1 term for each program.

For the estimation of $x_j(0)$ the total $N(0)$ was distributed in the n semesters keeping the proportion defined by the group of values $\tilde{x}_1(0), \tilde{x}_1(0), \cdots, \tilde{x}_n(0)$.

$$x_j(0) = \frac{\tilde{x}_j(0)}{\sum_{h=1}^{n} x_h(0)} N(0)$$

7) Real Number of Students: Using the report of active and blocked students it was calculated $N(k)$, the number of active students in term k for each program.

8) Initial Conditions for the School: Using the report of active and blocked students it was calculated $N_F(0)$, the number of active students in term 2009-1 of the School. This amount was distributed in the n semesters in such a way that there would be more students in the lower semesters than in the superior ones. It was used a geometric relation of the form

$$x_{j+1}(0) = m x_j(0) \qquad 0 \leq m \leq 1$$

The value of the initial condition for the first semester is obtained like this:

$$X_1(0) = N_F(0) * \frac{1-m}{1-m^n}$$

Where n is the number of semesters ($n = 10$) and the value of m was selected in 0.05.

9) Approval Rates for the School: The approval rates were estimated in such a way that the odds of approving a course in the superior semesters would be better than in the lower semesters. It was used a geometric relation on the repetition rates $((1 - \alpha_j(k))$. The repetition rates are constant trough time:

$$\alpha_{j+1}(k) = 1 - m_\alpha(1 - \alpha_j(k)) \qquad 0 \leq m_\alpha \leq 1 \qquad \forall k$$

The value selected for $\alpha_1(k)$ was $= 0.65$, which represents the rate of approval of the course 'Differential Calculus'[3]. The value of m_α was selected in $= 0.5$, which means that the repetition rate diminish at half of the semester.

10) Cancellation Rates for the School: The cancellation rate of the School has been rising since the 2009-1 term in a sustained fashion. For this reason the cancellation rate was modeled according to a related relation in function of time:

$$\beta_j(k) = \beta_0 + m_b k \quad 0 \leq \beta_0 \leq 1 \quad m > 0$$

By doing an adjustment to the straight line on the cancellation rates from the 2009-1 term to the 2012-3 term, the values that result are $\beta_0 = 0.11$ and $m_\beta = 0.03$

[3]The first course in figure 3

IV. Analysis of the reduced model

Using the reduced model the following questions of analysis are formulated:

P.1. How does the number of students in a program evolve if the enrollment is constant?

P.2. Which is the expected number of students of a program if the enrollment is constant?

P.3. What is the effect of the parameters of the model on the expected number of students?

P.4. How many students will graduate and how long would it take them?

P.5. What effect does the parameters of the model have on the number of students who graduate and the completion time?

These questions are approached in the following sections. Equation 6 represents a Linear Time Invariant Discret Dynamic System that can be analyzed using some well known concepts as ‘Transfer Function’, ‘Step response’, ‘Impulse Response’ and ‘Stationary state’.

A. Evolution of the number of students

To analyze the evolution of the number of students it is considered the case in which there are no student transfers ($t_j(k) = 0$) and the registrations are constant ($u(k) = u$). In other words, we conducted a ‘step response’ analysis using registrations as inputs.

If we consider $y_T(k)$, the total number of students, as the output of the system then a transfer function $F_T(z)$ can be defined as:

$$F_T(z) = \frac{Y_T(z)}{U(z)}$$

In [10] it is shown that $F_T(z)$ can be derived from equation 6:

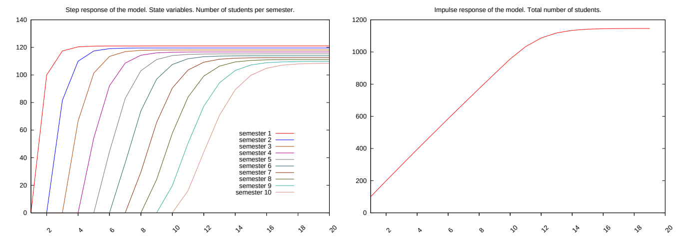
$$F_T(z) = \frac{1}{\gamma}\sum_{j=1}^{n}\frac{\gamma^j}{(z-\delta)^j}$$

The system has a single pole δ which appears n times in the transfer function. The stability of the system depends then on the value of δ. As α and β are values in the interval $[0, 1]$ therefore,

$$\begin{array}{lcl} 0 \leq & \gamma = (1-\beta)\alpha & \leq 1 \\ 0 \leq & \delta = 1-\gamma & \leq 1 \end{array}$$

Since δ is in the interval $[0, 1]$ the system is stable. In the limit case in which $\delta = 1$, the system would be marginally stable. For this to occur it is needed that $\gamma = 0$, which means that it is required that one the following conditions is meet:

- $\beta = 1$, which means that every student cancelled all of the courses.
- $\alpha = 0$, which means that no students approved their courses.



(a) Number of students per semester (b) Total number of students

Figure 7. Step response of the simplified model. $\alpha = 0.85$ $\beta = 0.03$ $\sigma = 0.01$

1) Example: To illustrate the behavior of the number of students a simulation was ran for 10 semesters ($n = 10$), with no transfers ($t_j(k) = 0$), with a constant number of 100 registrations per period ($u(k) = u = 100$) and with null initial conditions ($x_j(0) = 0$). The parameters of the model are fixed in $\alpha = 0.85$, $\beta = 0.03$ and $\sigma = 0.01$. This simulation represents the expected evolution of a new program.

Figure 7 shows the results; in figure 7a it is shown how the number of students enrolled in every semester evolves and figure 7b shows the evolution of the total number of students. When analyzing these figures it is found that:

- The system stabilizes after approximately 16 semesters.
- The value of the stationary value of each semester is different, being bigger the one of the inferior semesters. These differences are due to the students who do not reach the upper semesters because of desertion.
- The times of stabilization of each semester are different: the curves initiate in different time and show different inclinations, being smaller the ones from the upper semesters.

B. Expected number of students

To analyze the expected number of students, it is calculated the total number of students when many terms have passed, without transfers ($t_j(k) = 0$) and with constant registrations ($u_j(k) = u$). In other words, the value of the variable $y_T(k)$ in stationary state is evaluated (see [10])

$$y_{ee} = \lim_{k\to\infty} y_T(k)$$

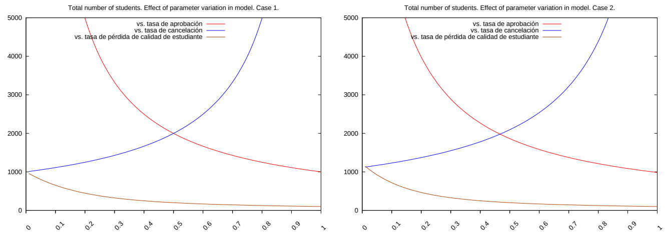
$$y_{ee} = \frac{u}{\sigma}\left(1 - \left(\frac{(1-\beta)\alpha(1-\sigma)}{\sigma + (1-\beta)\alpha(1-\sigma)}\right)^n\right)$$

1) Case without desertion: If there are no desertion ($\sigma = 0$), the expected value of the number of students must be calculated as:

$$y_{ee} = \frac{nu}{\gamma} = \frac{nu}{(1-\beta)\alpha}$$

2) Case of reference: A special case is considered: If there is no cancellations, nor desertion and every student approve all the subjects, the number of students is reduced to:

$$y_{ref} = y_{ee} = nu$$



(a) Case 1 (b) Case 2

Figure 8. Effect of the parameters on the expected number of students

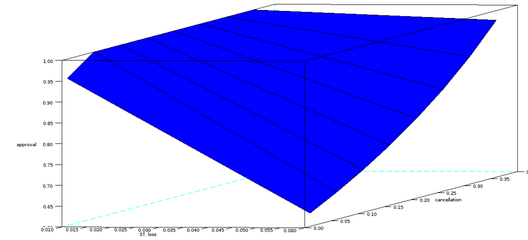


Figure 9. Reference value surface

C. Effect of the parameters on the expected number of students

To evaluate the effect of the parameters of the model on the expected number of students, it was considered a 10 semester program ($n = 10$) in which 100 students registered each period $u = 100$. With these considerations, the reference value is 1.000 students ($y_{ref} = nu = 1000$).

Using this case as a base for comparisnos, the following experiments were carried out:

- Case 1: There is a variation in the parameters of the model α, β and σ one at a time. The variation was done in fact from 0 to 1 and the expected number of students is calculated: In each case, when a parameter remains constant, it takes the following values:
$$\alpha = 0 \qquad \beta = 0 \qquad \sigma = 0$$
- Case 2: There is a variation in the parameters of the model α, β and σ one at a time. The variation was done in fact from 0 to 1 and the expected number of students is calculated. In each case, when a parameter remains constant, it takes the following values:
$$\alpha = 0.83 \qquad \beta = 0.01 \qquad \sigma = 0.03$$

The results of the experiments are shown in figure 8. When analyzing this figures it is evident that:

- The total number of students is very sensitive to the cancellation and approval rates.
- The sensitivity of the total number of students to the cancellation and approval rates increases in Case 1, which means that increases when desertion rate decreases.
- The current cancellation rate is near to 0.2. For these values, the cancellation might be causing an increase of up to 20% in the total number of students.

1) Expected value and Reference value: Figure 8 shows that the effect on the expected number of students of some parameters is increasing and in others is decreasing.This fact generates a question for analysis:

- What combination of parameters makes the expected number of students be equal to the reference number? Or – which is the same –, under what circumstances it is satisfied:
$$\frac{u}{\sigma}\left(1 - \left(\frac{(1-\beta)\alpha(1-\sigma)}{\sigma + (1-\beta)\alpha(1-\sigma)}\right)^n\right) = nu$$

Fort the case in which $n = 10$, the combination of parameters is.

$$\alpha = \frac{\sigma/(1-\beta)}{(1-10\sigma)^{-0.1} - 1} \tag{7}$$

Figure 9 shows the tridimensional area which is generated from (7).

D. Graduation times

Not all the students that are enrolled in an academic program graduate; the students who graduated do it in a different number of semesters. To explore the times of graduation of the admitted students in period m (who entered the program in the term $m+1$) the graduation series $G_m(k)$ is defined as:

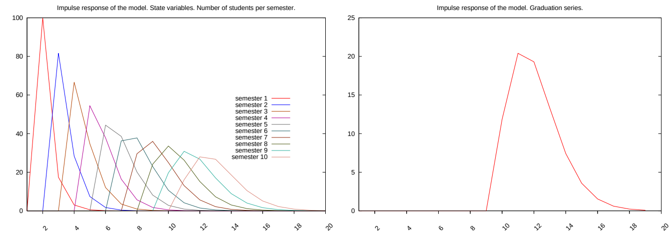
$$G_m(k) = g(k)|_{impulso} \tag{8}$$

The subindex $impulse$ denotes the 'Impulse Response' analysis conditions:

- Null initial conditions: $x_j(0) = 0$
- Null student transfers: $t_j(k) = 0$
- The series of registration represents a discrete impulse in period m:
$$u(k) = \begin{cases} u & \text{si } k = m \\ 0 & \text{si } k \neq m \end{cases} \tag{9}$$

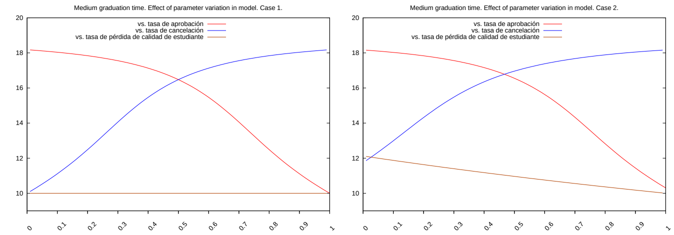
1) Indicators of graduation time: By using the series $G_m(k)$) several indicators on the graduation process can be defined:

- Total number of students who graduate GT_m: corresponds to the sum of all the terms of the series:
$$GT_m = \sum_{k=0}^{\infty} G_m(k)$$
- Average time of graduation $\overline{G_m}$: the time between the enrollment to the program (which happens in term $m+1$) and the time k is $(k-(m+1)) = (k-m-1)$; therefore, the average time of graduation is calculated as:
$$\overline{G_m} = \frac{\sum_{k=0}^{\infty}(k-m-1))G_m(k)}{GT_m}$$
- Variance in the graduation time: calculated as
$$GV_m = \mathrm{Var}[G_m(k)] = \frac{\sum_{k=0}^{\infty} G_m(k)\left(G_m(k) - \overline{G_m}\right)^2}{GT_m}$$



(a) Number of students per semester (b) Total number of students

Figure 10. Impulse response of the simplified model. $\alpha = 0.85$ $\beta = 0.03$ $\sigma = 0.01$



(a) Case 1 (b) Case 2

Figure 11. Effect of the parameters on the graduation times

If the graduation time is considered for a single cohort that enters in the initial term ($m = 0$), then the expressions for the graduation series and the average time of graduation are:

$$
\begin{aligned}
G_0(k) &= \begin{cases} \frac{\delta^{k-n}\gamma^n(k-1)!}{(n-1)!(k-n)!} & \text{si } k \geq n \\ 0 & \text{si } k < n \end{cases} \\
\overline{G_0} &= \sum_{k=n}^{\infty} \frac{\delta^{k-n}\gamma^n k!}{(n-1)!(k-n)!} \\
\gamma &= (1-\beta)\alpha(1-\sigma) \\
\delta &= \beta(1-\sigma) + (1-\beta)(1-\alpha)(1-\sigma)
\end{aligned} \tag{10}
$$

2) Example: To ilustrate the behavior of the number of students a simulatrion was ran for 10 semesters ($n = 10$), with no transfers ($t_j(k) = 0$), with a constant number of 100 registration in the initial term ($u(0) = u = 100$) and with null initial conditions ($x_j(0) = 0$). The parameters of the model are fixed in $\alpha = 0.85$, $\beta = 0.03$ and $\sigma = 0.01$. This simulation represents the evolution of a single cohort.

Figure 10 shows the results of the simulation. In figure 10a it is shown how evolves the number of students enrolled in each semester and figure 10b shows the evolution of the total number of students. For this example, the indicators of the graduation time take the following values:

$$
\begin{aligned}
GT_m &= 88.5 \\
\overline{G_m} &= 12.1 \\
GV_m &= 2.5
\end{aligned}
$$

E. Effect of the parameters on the graduation times

To evaluate the effect of the parameters of the model on the expected number of students, it is considered a program of 10 semester ($n = 10$) to which every term enter 100 students $u = 100$. With these conditions, the value of reference is 1.000 students ($y_{ref} = nu = 1000$).

With this base program the following experiments were carried out:

- Case 1: There is a variation in the parameters of the model α, β and σ one at a time. The variation was done from 0 to 1 and the average time for graduation is calculated. In each case, when a parameter remains constant, it takes the following values:

$$\alpha = 0 \qquad \beta = 0 \qquad \sigma = 0$$

- Case 2: There is a variation in the parameters of the model α, β and σ one at a time. The variation was done from 0 to 1 and the average time for graduation is calculated. In each case, when a parameter remains constant, it takes the following values:

$$\alpha = 0.83 \qquad \beta = 0.01 \qquad \sigma = 0.03$$

The results of the experiments are shown in figure 11. By analyzing these figures it is found that:

- The average time of graduation is very sensitive to the cancellation and approval rates.
- The average time of graduation is not very sensitive to the desertion rate.
- The sensitivity of the total number of students to the cancellation and approval rates is similar in the two cases.
- The current cancellation rate is close to 0.2. For these values, the cancellation rate might be causing and increase of up to two semesters in the average graduation times.

V. Analysis of the increase of students at the School of Engineering

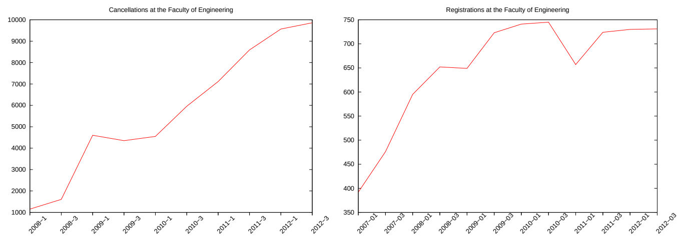
To explain the increase in the number of students at the School of Engineering shown in Figure 2b, the following possible causes are formulated:

C.1. **The increase in the number of freshmen:** Figure 12b shows the evolution of the students who entered the School by an admission exam since 2007. There is a significant growth in 2009, and an exceptional decrease in the 2011–1 period.

C.2. **The increase in the cancellation rate:** Figure 12a shows the evolution of the number of cancellations of subjects previously inscribed by students of the School. The continued growth it is explained by the normative changes of 2008.

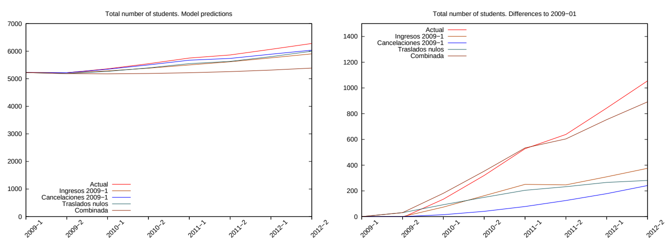
C.3. **The increase of the transfers and double degree enrollments:** Double degree programs is a relatively new phenomenon. It was created in 2008 and started to be effective in 2010.

C.4. **The combined effect of the pervious causes**.



(a) Cancellations (b) Registrations

Figure 12. Evolution of two feasible causes of the total number of students increase



(a) Number of students (b) Differences to 2009-01

Figure 13. Analysis of feasible causes of the total number of students increase

Table I
CONTRIBUTION OF EACH POTENTIAL CAUSE TO THE EXPLANATION OF THE INCREASE IN THE NUMBER OF STUDENTS AT THE SCHOOL.

Potential Causes	**Percentage of explanation**
Registrations	35.60%
Cancellations	22.91%
Transfers	26.64%
Combination	84.54%

To explore the impact of the possible causes, four simulation experiments were designed. The four experiments take as a base the simulation conditions used to validate the model with the School data. From that base condition, some modifications on the simulation conditions were made for each experiment:

Ex.1. The registration of freshmen was kept constant in the value corresponding to the 2009–1 term.
Ex.2. The cancellation rate was kept constant in the value corresponding to the 2009–1 term.
Ex.3. It was considered null transfers during the simulation.
Ex.4. The three previous conditions were simultaneously considered.

The results of the simulation are shown in figure 13. Figure 13a shows the prediction of the total number of students for each of the experiments, while figure 13b shows the difference of that same number compared to the total number of students in the 2009–1 period. It is evident how each of the potential causes partially explains the real growth. Chart I shows the contribution of each of the potential causes to the explanation of the increase in the number of students at the School. The model achieves a partial explanation of 84.54% of the increase.

VI. CONCLUSIONS

The following conclusions refer to the model discused in this paper:

C.1. The mathematical model proposed can model properly the cancellation of courses, the repetition, approval, registration, transfers (and double degree programs), desertion and graduation in the context of the School of Engineering of the National University of Colombia.
C.2. When analyzing the validation of the model it is evident that this one adjust better for: a) the total number of students rather than the number of students who graduate and b) the School as a whole rather than individual programs.
C.3. From historical data it is possible to estimate the specific parameters of the model for a specific program. This task, however, is not expedited, given the very structure of the institutional information available.
C.4. The simplified model allows to do a detailed mathematical analysis of both the total number of students and of the time of graduation.
C.5. The increase registered from the 2009–1 term in the cancellation rate has a clear impact on both the total number of students in the program and on the time of graduation. The increases registered might be generating a two academic semester's increase in the graduation times and of the 20% of the total number of students.
C.6. The increase of the number of students at the School of Engineering is mainly explained by the increase in the students who entered by admission test. The cancellations and transfers (added to the double degree program entries) were also an important cause of the mentioned increase.

We must emphasize that the validaty of the model has been analyzed just in the context of the School of Engineering of the National University of Colombia. However, the model is enough general to be adapted for other contexts.

Using this model we are able to predict the number of students in every level of a program. This prediction can be used to quantify the resources needed to attend the teaching demand. Actually, authors are working in a teaching demand vs teaching capacity model for the same context.

ACKNOWLEDGMENT

The authors want to thank the reviewers for their recommendations and suggestions which have helped improve the quality of this work. This work has been possible thanks to the support provided by Universidad Nacional de Colombia.

REFERENCES

[1] UniSaarland, "Methodology of the calculation of student capacities in Germany. An Introduction," Saarland University, Tech. Rep., 2013.
[2] CSU, "Student flow analysis: CSU Student Progress Toward Graduation Report to the California State University As part of the Making Opportunity Affordable Planning Year," Institute for higher education leadership and policy. California State University, Tech. Rep., 2009.

[3] C. Rosa and E. Pereira, “A Metapopulation Model for the Study of the Evolution of the Number of Students in a Teaching Institution,” in *2013 8th Iberian Conference on Information Systems and Technologies (CISTI)*, 2013.

[4] UNC, “Acuerdo 033 de 2008 del Consejo Superior Universitario,” http://www.legal.unal.edu.co/sisjurun/normas/Norma1.jsp?i=34245, 2007.

[5] E. Caro, C. González, and J. M. Mira, “Student academic performance stochastic simulator based on the Monte Carlo method,” *Computers & Education*, vol. 76, pp. 42–54, 2014.

[6] B. H. Wells, H. A. Sanchez, and J. M. Attridge, “Modeling Student Interest in Science, Technology, Engineering and Mathematics,” in *Meeting the Growing Demand for Engineers and Their Educators 2010-2020 International Summit, 2007 IEEE*, 2007.

[7] J. J. Lin, K. J. Reid, and P. Imbrie, “Work In Progress - Predicting Retention in Engineering using an Expanded Scale of Affective Characteristics from Incoming Students,” in *39th ASEE/IEEE Frontiers in Education Conference*, 2009.

[8] A. Pérez-Gama, G. Hoyos, L. Parra-Espitia, M. Ortegón, L. G. Rozo-Pardo, and B. Perez-Gutierrez, “Education Software Architecture: Facing Student Desertion in Colombia Higher Education with an Intelligent Knowledge Based Coaching System,” in *ANDESCON 2010*, 2010.

[9] R. Alkhasawneh and R. Hobson, “Modeling Student Retention in Science and Engineering Disciplines Using Neural Networks,” in *2011 IEEE Global Engineering Education Conference (EDUCON)*, 2011.

[10] O. Duarte, “Un modelo del tránsito de los estudiantes a través del plan de estudios,” Universidad Nacional de Colombia. Facultad de Ingeniería. Vicedecanatura Académica. Available in http://www.bdigital.unal.edu.co/11827/, Tech. Rep., 2013.