Relationships between student throughput variables and properties

Many different models have been designed to describe the plethora of factors that influence student throughput and success and how these factors affect throughput system variables and properties. System variables include headcounts (H) and successful credits (S) of throughput systems; some examples of system properties are the percentage of the new student intake graduating annually, and the average number of years to graduate or to drop out of a degree. However, no past study has defined the analytical relationships between these variables and properties from a process perspective – which was the purpose of this study. Three simple analytical equations were derived for 4-year degrees, and then geometrically interpreted. The behaviour of a simplified throughput system can be described by the position of a point in the admissible region of the H-S plane, with each point relating to a specific set of system properties. The successful credits ratio (S/H) is shown to be the ideal process efficiency ratio for throughput systems. The results were also extended to degrees of shorter duration. The behaviour of real throughput systems is broadly found to be similar to the behaviour of simplified throughput systems. In this study, only the mathematical foundations for the general relationships between throughput properties and throughput variables for a degree were established. The way in which this mathematical basis finds application in practice is illustrated for a few selected cases only, because of the specific focus of this paper.

Introduction

Understanding the complexities of the throughput patterns of students enrolled for degree studies at universities has always been a major challenge. Numerous models are reported in the literature to describe the plethora of factors that influence student throughput and success and how these factors affect throughput system variables and properties. A review of the current literature in this regard can be found in recent studies. From a process perspective, however, no study could be found that defines the analytical relationships between the most important variables and properties of a student throughput system; this absence is the reason for the current study. The throughput variables included in this study were the headcount number of students, the total number of successful module credits earned by students and the full-time equivalent (FTE) value for students enrolled for the degree. The annual intake of new students was also a special throughput variable considered here. Throughput properties considered were the percentage of new entrants graduating annually as well as the average time to graduate or to drop out of a degree. The word ‘degree’ in this paper is used in a generic sense and also refers to a diploma or certificate. In this paper, a simplified cohort survival model for a 4-year degree will be used to establish analytical formulae for the key system variables as functions of the system properties and the annual intake of new students. A geometrical interpretation of the analytical equations will also be given to show exactly how all the system variables and system properties are connected to one another simultaneously. The results developed for a 4-year degree will then be extended to degrees of shorter duration. This paper only establishes the mathematical basis for the general relationships between throughput properties and throughput variables for a degree. Although the application of the results to real throughput systems will be discussed, the way in which this mathematical basis finds application in practice will be illustrated for a few selected cases only. The theory as presented is not a model of South African higher education, but can be used to assess throughput process efficiency within the local higher education system as well as in other systems. The many interventions that can be undertaken by universities to improve the efficiency of the throughput process are equally important but are excluded from the scope of this paper.

Simplified cohort survival model calculations

The background calculations in this paper are structured in terms of a standard cohort survival model for a 4-year degree with a maximum time for completion assumed to be 8 years. The 8 years cut-off is necessary to ensure that the analytical expressions that will be derived are of a simple form. Treating the number of cut-off years as a variable increases the number of parameters in the model without necessarily contributing to a better understanding of throughput systems. In a simplified student throughput system, the cohort size remains constant, and the student graduation and dropout patterns for cohorts repeat year after year. These characteristics give rise to stationary throughput patterns that have also been referred to as equilibrium throughput patterns. The standard cohort survival model used, together with the simplifying assumptions made in this paper, created the simplest model possible for deducing analytical expressions for the various throughput variables in terms of throughput properties. Cohort survival models in general produce numbers as outcomes with very little chance of discovering the analytical relationships that exist between the said entities. The details of the cohort calculations documented in this paragraph as new research results are important in order to understand the full impact of this work, but may be skipped by readers not necessarily interested in such detail. A summary is provided in the next paragraph.

In the simplified cohort survival model, the student throughput history for a 4-year degree will be reflected by the simultaneous presence of a set of eight cohorts with cohort 1 being the youngest cohort in the year under observation. Each cohort has the same throughput profile characterised by the percentage of the cohort graduating or dropping out from the system at the end of the year. Hence, if \( G_i \) is the percentage of the annual intake \( N \) of
The headcount $H_D$ of students who will eventually graduate can be derived through the cohort approach as:

$$H_D = (1D_1 + 2D_2 + \ldots + 7D_7 + 8D_8)N = (1-G)KN. \quad \text{Equation 5}$$

The total headcount $H$ for the throughput system is therefore defined in terms of the three independent system properties $G$, $J$ and $K$ as:

$$H = [GJ + (1-G)KN] = [K + G(J-K)N]. \quad \text{Equation 6}$$

This formula resembles a similar formula that was derived in a different way by Breneman for a production function for PhDs, as reported in Hopkins’. To calculate the total successful module credits earned annually by students who will eventually graduate (a maximum of 1 credit per student per year), the following is considered: a student who graduates in 4 years would earn 4/4 successful module credits per year; a student who graduates in 5 years would earn an average of 4/5 successful module credits per year; etc. The total number of successful module credits $S_G$ earned annually by students who will eventually graduate is then given by:

$$S_G = \frac{4G_4 + 5G_5 + 6G_6 + 7G_7 + 8G_8}{G}. \quad \text{Equation 2}$$

and the average number of years $J$ studied by students dropping out at the end of the year is derived similarly as:

$$J = \frac{(4G_4 + 5G_5 + 6G_6 + 7G_7 + 8G_8)}{G}. \quad \text{Equation 3}$$

Similarly, the average number of years studied by students dropping out at the end of the year under observation is denoted by $J$. The value of $J$ ranges between 4 and 8, and the value of $K$ ranges between 1 and 8. The three quantities $G$, $J$ and $K$ are independent of one another. They also describe the main characteristics of the student throughput process and are therefore referred to as throughput system properties. The characteristics of a throughput system can therefore be described in terms of all the possible combinations of $G$, $J$ and $K$. System variables are used to describe bulk system quantities that only change if the properties of the throughput system change. These system variables are the headcount $H$ of the throughput system, the successful module credits $S$ earned annually by students, and the FTE value for students enrolled for the degree $V$. It has already been shown in Equations 6 and 9 that the system variables $H$ and $S$ depend on the system properties $G$, $J$ and $K$ on the annual intake of new students $N$. Student headcount $H$, which in the South African higher education system refers to an unduplicated count of students irrespective of the academic course load of the student, depends on $G$, $J$ and $K$ as follows:

$$H = [GJ + (1-G)KN] = [K + G(J-K)N]. \quad \text{Equation 6}$$

The perfect throughput system in which all students graduate in minimum time, has a size equal to $H=4N$ when $G=100$% and $J=4$. The influence of $K$ diminishes as $G$ approaches 100%. The total successful module credits $S$ earned annually by students enrolled for a 4-year degree (a maximum credit of 1 per student per year) depends on $G$ and $K$ as follows:

$$S = [4G + 0.25(1-G)KN]. \quad \text{Equation 9}$$

The number of successful module credits $S$ earned annually by students for the perfect throughput system is equal to $S=4N$ when $G=100$%. Again, the influence of $K$ diminishes as $G$ approaches 100%. Apart from the successful credits ($S$) earned annually by students, the credits assigned to modules not successfully passed by students in the same year will be referred to here as failed credits ($F$). Furthermore, some students often take fewer modules than required by a full academic load with module credits therefore adding up to less than 1. In this paper, the balance of module credits not attempted by students in a particular year is referred to as unutilised credits ($U$). Using the fact that each of the headcount students can at most generate 1 credit per year, it clearly follows that:

$$S + F + U = H \text{ or } F + U = H - S. \quad \text{Equation 10}$$

The FTE value for students enrolled for the degree $V$ as a throughput variable, which in the South African higher education system depends on
the successful credits as well as the failed credits for students enrolled for the degree, is defined by:

$$V = S + F = LS + (1 - L)H,$$  
Equation 11

where the fourth system property $$L = U/(U + F)$$ defines the balance of credits between $$U$$ and $$F$$. This fundamental relationship between $$V$$, $$S$$ and $$H$$ implies that $$V$$ will range between $$S$$, when $$L = 1$$ with no failed credits ($$F$$) present, and $$H$$, when $$L = 0$$ with no unutilised credits ($$U$$) present. In a perfect throughput system, $$H = 4N = S$$ with the FTE value $$V$$ also being equal to $$4N$$. In the South African higher education system, the biggest part of the funding of universities, as well as the provision of building facilities, mainly depends on $$V$$. The FTE value is also a direct measure of the actual academic load on students in a particular year and is therefore also generally used by universities as a basis for the provision of lecturing staff.

### Relationship between $$H$$ and $$S$$ in the throughput system

The system variables $$H$$ and $$S$$ defined above are not independent but are each dependent on the same set of $$G$$, $$J$$, and $$K$$ values for the student throughput configuration under consideration. This connection is mathematically defined by combining Equations 6 and 9 under the assumption that $$K$$ be treated as a parameter. This definition establishes specific relationships between $$H$$ (as well as $$S$$) and the two system properties $$G$$, $$J$$, and $$K$$. These relationships can be made visible through graphs in the $$J$$-$$G$$ plane, which unfortunately produces rather complex patterns of $$H$$ and $$S$$. This complexity can be avoided by analysing these relationships in the $$H$$-$$S$$ plane, as shown in Figure 1. Such an analysis reveals that only certain combinations of $$H$$ and $$S$$ can be realised, namely those included in the so-called admissible region of the $$H$$-$$S$$ plane. The admissible region is a triangle, $$WXZ$$, bounded by the line $$WX$$ representing $$J = 4$$, by the line $$WZ$$ representing $$J = 8$$, as well as by the horizontal lines of $$G = 0\%$$ and $$G = 100\%$$. A second vertical axis on the right has been added to show the corresponding values of $$G$$. The parameter $$K$$ has been set equal to $$4$$ by way of example with the latter value corresponding to a specific system property. The value of $$K$$ also defines the position of the pivot $$W$$ of lines of constant $$J$$ and their intersection with the line $$G = 0\%$$. In the admissible region $$WXZ$$, pairs of admissible $$H$$ and $$S$$ values, such as the throughput configuration $$W$$ with $$H = 5N$$ and $$S = 3N$$, would always be connected to lines of constant $$J$$ values (in this case $$J = 5.5$$) and lines of constant $$G$$ values (in this case $$G = 67\%$$). The perfect throughput system is located at $$X$$ now defined by $$H = 4N = S$$ and is produced by the intersection of the lines $$G = 100\%$$ and $$J = 4$$. Only within the admissible region defined by a given value of the parameter $$K$$, will each combination of $$G$$ and $$J$$ correspond to a unique combination of $$H$$ and $$S$$, and vice versa.

Of particular importance in this analysis is the successful credits ratio for the simplified throughput system defined in this paper by the ratio $$S/H$$. This ratio which can be regarded as the ideal process efficiency ratio compares the output of the successful credits $$S$$ produced during the year under consideration to the input of the total credits $$H$$ available for that year, both $$S$$ and $$H$$ depend on the same set of system properties $$G$$, $$J$$, and $$K$$. The successful credits ratio is considered to be ideal in the sense that it will be shown to apply to simplified as well as real throughput systems. The successful credits ratio as a single number produces a simultaneous account of the combined efficiency status of the throughput system characterised by a specific set of system properties $$G$$, $$J$$, and $$K$$. The successful credits ratio for the throughput system $$Y$$ is given by $$S/H = 60\%$$. As the efficiency of the throughput system increases, the lines of constant $$S/H$$ move towards $$X$$. The successful credits ratio $$S/H$$ will be equal to $$100\%$$ for the perfect throughput system when $$G = 100\%$$ and $$J = 4$$. This scenario is true for all degrees irrespective of the duration of the degree.

Furthermore, lines of constant FTE values are defined by $$V/N = \text{constant}$$. If $$L = 0$$, the lines $$V/N = \text{constant}$$ are vertical lines, and if $$L = 1$$, these lines are horizontal; all of these lines are also parallel to one another. The line $$V = 4N$$ (with $$L = 0.5$$) is also shown in Figure 1 as a dotted line passing through $$X$$. The fact that $$Y$$ lies on the line $$V = 4N$$ implies that the $$V$$ value of the throughput system $$Y$$ is $$4N$$. Above the line $$V = 4N$$, the FTE values $$V$$ are larger than $$4N$$, and below the line, the values are smaller than $$4N$$. With much of the attention presumably focused on the migration of the system $$Y$$ along the line $$V = 4N$$, it would appear that a value of $$K$$ equal to 4 would conveniently be required to restrict migrations of $$Y$$ diverging too far to the left. This situation implies that $$K$$ as a system property would have to be managed in such a way as to remain at a value equal to the minimum time of completion of the degree. The geometrical interpretation of Equations 6, 9 and 11 is also shown in Figure 1, again illustrating the interdependence of throughput variables and properties. As the throughput system $$Y$$ migrates within the triangular admissible domain, $$Y$$ carries along with it the constant of $$H/N$$, $$S/N$$, $$V/N$$, $$G$$, $$J$$, and $$L$$, simultaneously showing the relationships between these quantities. The value of $$K$$ defining the pivot $$W$$ of the triangle, can be read from the $$H$$ axis with various possible positions of the pivot as shown by the square markers. The value of $$J$$ can also be read from the $$H$$ axis at the top, and the value of $$G$$ can be read on the right-hand axis. Constant successful credits ratio lines $$S/H = \text{constant}$$ all pass through the fictitious origin of the $$H$$-$$S$$ plane. In summary, the behaviour of a simplified throughput system can be described by the position of a point in the admissible region of the $$H$$-$$S$$ plane, each point relating to a specific set of system properties. The admissible region is bounded by lines of maximum and minimum values of $$G$$ and $$K$$ with $$G$$ preferably managed to be equal to the minimum time for completion of the degree. The migration of the throughput system in the $$H$$-$$S$$ plane means changing the system properties and allowing enough time for the system to establish equilibrium in its throughput patterns. The successful credits ratio for the system is equal to $$S/H$$ and its FTE value is equal to $$(S+H)/2$$ for $$L = 0.5$$. For a given value of $$N$$, the perfect throughput system is found at $$H = 4N = S$$ with $$S/H = 100\%$$.

### Throughput systems for 1- to 4-year degrees

Here the simplified model for 4-year degrees is extended to degrees of shorter duration. In the case of the 4-year degree, it was assumed that the minimum time of 4 years to be taken by students to graduate should be limited to a maximum of 8 years. It is suggested that the same rule be applied for degrees of shorter duration, the reason being the simplification of the analytical expressions to be derived. In the case of degrees of shorter duration, the relevant simplified cohort survival model can be used to derive simple expressions for $$H$$ and $$S$$, similar to Equations 6 and 9. The assumption of the maximum time for completion of the degree being double the minimum time for completion, leads to strikingly similar analytical expressions, each being a function of the minimum time for completion of the degree. Therefore, with $$M$$ the minimum time
for completion of a degree in general, \( L \) will then range between 2 and 2M, whereas \( K \) would range between 1 and 2M. The headcount \( H \) of the student throughput system is given by \( H = \{K+G/(J-K)\}N \), whereas the successful module credits \( S \) earned annually by students would then be given by \( S = [MG + 0.25(1-G)J]N \).

**Application of the findings to a few selected topics**

The main purpose of this paper is to establish the mathematical basis for the general relationships between throughput properties and throughput variables for a degree. However, a few simple applications of the equations derived in this paper will now be given for degrees offered in the South African higher education system. More complex applications to issues such as the conversion of 3-year degrees into 4-year degrees for students who are expected to find it difficult to complete the 3-year degree in minimum time,\(^5\) will be discussed in a follow-up paper.

**Clarity on process efficiency measurements**

In the South African higher education system, the success rate \( S/(S+F) \) that can be written as \( S/(S+F) = (S/H)/(1-L(1-S/H)) \) is used as a measure of efficiency instead of the successful credits ratio \( S/H \) as defined in this paper. The success rate will clearly measure the process efficiency of degrees differently because of its dependence on \( L \). Degrees with \( L \) values close to 1 and a success rate therefore equal to 1, will seem to be more efficient than degrees with \( L \) values close to 0 and a success rate therefore equal to approximately \( S/H \). This difference in efficiency measurement becomes very pronounced in the case of degrees with relatively low successful credits ratios, such as \( S/H = 50\% \). The use of the success rate should therefore be discontinued in favour of the use of the successful credits ratio, \( S/H \).

Another ratio that is often used in the South African higher education system as a measure for process efficiency is the FTE to headcount ratio \( V/H \). With \( V \) equal to \( LS+(1-L)H \), it is clear that this ratio would likewise produce different measurements of process efficiency as a result of its dependence on \( L \); its use should therefore also be discontinued.

The graduate output to size ratio \( GN/H \) is known in the South African higher education system as the graduation rate, and its application to real throughput systems is widely contested because of fluctuations in both the values of \( G \) and the intake \( N \) of new students.\(^6\) Despite this disadvantage in the case of real systems, it may still be useful in simplified systems as an absolute output to size ratio, focusing on the process efficiency of the teaching learning process irrespective of the duration of the degree. However, a 4-year degree with \( GN/H \) being equal to 0.25 at best, would then from a process perspective appear to be less efficient than a 3-year degree with its ratio \( GN/H \) equal to 0.33 at best. This outcome would certainly restrict the use of this ratio as a process efficiency measurement for degrees in general. The general perception that \( GN/H \) is an unrelated or even a more comprehensive measurement of process efficiency is also incorrect. The fact is that the two measurements \( S/H \) and \( GN/H \) are related to one another through Equation 9 with \( (GN/H) = (S/H)/4 = 0.25 \) for the perfect throughput system. It is noted that lines of \( GN/H \) = constant all pass through the fictitious origin of the H-G plane; in particular, the line \( GN/H = 0.133 \) passes through \( Y \) in Figure 1.

**Inequities in the funding framework for South African universities**

The biggest part of the funding framework for South African universities is based on the sum total of the FTE values \( V = LS+(1-L)H \) for each university degree. Degrees with \( L \) values close to 1 will have FTE values very close to \( S \) and therefore a funding base that will be largely performance driven through the output variable \( S \). However, degrees with \( L \) values close to 0 will have FTE values very close to \( H \) and therefore a funding base that will be largely input driven through the input variable \( H \). This difference is considerable, especially in the case of degrees in which the successful credits ratio is \( S/H = 50\% \). In such a case, the \( L=0 \) degree will receive double the funding received by the \( L=1 \) degree. The question is whether two degrees with exactly the same successful credits ratio \( S/H \) should be funded at such disparate levels especially if no specific reason can be identified to justify the existence of different values of \( L \)? It therefore seems reasonable that \( L \) should be assumed to be 0.5 for all degrees and that the quantity \( LS+(1-L)H = (S+H)/2 \) be used as a more appropriate basis for the funding of degrees offered by South African universities. This approach would result in all degrees being funded on the basis of their average of \( S \) and \( H \), which, however, does not affect the actual \( L \) value for the degree, although one may eventually find a tendency amongst faculties to manage the \( L \) values of their degrees towards \( L=0.5 \). Consideration could perhaps also be given to change the definition of the FTE value to be based on the average of \( S \) and \( H \). It is noted that, whereas \( H \) can be regarded as the nominal size of the degree relating to an unduplicated count of student names on a list, the quantity \( S \) can be regarded as the credit earning size of the degree. The redefined FTE values \( (S+H)/2 \) could then be regarded as the effective size of the degree. Again, determining the full implications for the funding of South African higher education should, because of the complexity of the topic, require much more research and should rather be pursued outside of the scope of this introductory paper.

**Enrolment management**

The size of the South African higher education system is currently regulated by government-approved enrolment plans for universities with quotas for both \( H \) and \( V \) set for each university for each year. A minor relaxation of these constraints has recently been proposed,\(^7\) which will not address the difficulties highlighted below. The enrolment plans also call for meaningful annual improvement of the success rates and graduation rates for each university, thus signalling that higher successful credits ratios \( S/H \) need to be achieved for each university. These enrolment plans have three unintended consequences. Firstly, by imposing the FTE value constraint on throughput systems, universities would have to manage the \( H \) part of this constraint at the beginning of each academic year according to Equation 11 but would only know the value of the \( S \) part at the end of the year as a consequence of successful teaching outcomes. Such a constraint is very difficult to manage in practice. Secondly, when written in the form \( V/H = [1-L(1-S/H)] \), Equation 11 states that a student throughput system in which both \( V \) and \( H \) are constant, or even directly proportional to one another, can only produce the same but not higher successful credits ratios \( S/H \) as required by the enrolment plans. This dilemma can apparently be resolved by only retaining the constraint on \( V \), which is the more important constraint relating to the funding of the system. However, this again leaves the university with a constraint which is very difficult to manage in practice. Thirdly, imposing headcount quotas on the throughput system has led to enrolment management practices at South African universities which amount to registering returning students first and then using the new student intake \( N \) to make up for the shortfall. This unfortunate way of setting the size of the new student intake \( N \) introduces awkward fluctuations into the student throughput patterns during subsequent years, which in turn undermines proper planning with regard to the provision of facilities and lecturing staff. These three unintended consequences of enrolment planning within the South African higher education system, also point to the need for further research to be undertaken to resolve these difficulties.

**Real student throughput systems**

A better understanding of real throughput systems should follow directly from a study of simplified systems. Simplified systems are not theoretical constructs but in fact special cases of real throughput systems. The behaviour of a simplified throughput system can be described by the position of a point in the admissible region of the \( H-S \) plane, with each point relating to a specific set of system properties. The successful credits ratio of the throughput system is equal to \( S/H \) with the redefined FTE values equal to \( (S+H)/2 \). The same behaviour, however, is observed from the data for real throughput systems. This is explained by the fact that \( H \) can be calculated independently of Equation 6 by simply adding together the actual unduplicated number of students enrolled for a
In this paper, the mathematical foundation for the general relationships between throughput properties and throughput variables for a degree has been established using a simplified or equilibrium cohort survival approach. The simplified model assumes a constant annual intake of new students and that throughput system properties such as the graduation and dropout patterns for each cohort also remain the same. Throughput properties include the percentage \( G \) of the annual intake of new students graduating annually as well as the average number of years \( J \) to graduate and the average number of years \( K \) to drop out of a degree. The balance \( L \) between the unutilised and failed module credits within the throughput system is the fourth system property required. Three analytical formulae have been derived for important system variables, such as the headcounts \( H \) and total successful module credits \( S \) of the throughput system, both of which depend on the system properties \( G, J \) and \( K \) as well as on the annual intake of new students \( N \). The FTE value \( V \) has been expressed in terms of \( S, H \) and \( L \).

In this paper, it has been demonstrated that the system variables \( H \) and \( S \) are not independent but are each dependent on the same set of \( G, J \) and \( K \) values for the simplified student throughput configuration under consideration. Furthermore, only certain \( H \) and \( S \) values can simultaneously be realised – namely those in the admissible region of the \( H-S \) plane. The shape of the triangle is determined by the value of \( K \), which should preferably be managed to be equal to the minimum time for completion of the degree. Furthermore, within this triangle, the relationships between throughput variables and throughput properties become visible through suitable geometrical constructions. In essence, the behaviour of a simplified throughput system can be described by the position of a point in the admissible region of the \( H-S \) plane, with each point relating to a specific set of system properties. The migration of the throughput system in the \( H-S \) plane means changing the system properties and allowing enough time for the system to establish equilibrium in its throughput patterns. The successful credits ratio \( S/H \) produces a simultaneous account of the combined efficiency status of the three system properties \( G, J \) and \( K \). More importantly, the FTE number of students is given by \( (S+H)/2 \) for \( L=0.5 \), and for a given value of \( N \), the perfect throughput system for a 4-year degree is located at \( H=4N=S \) with \( S/N=100\% \). This paper provides indications on how the simplified model for a 4-year student throughput system can be changed to apply to degrees of different duration.

**References**

2. Prinsloo P. Modelling throughput at Unisa: The key to successful implementation of ODL. Pretoria: University of South Africa; 2009. Available from: http://hdl.handle.net/10500/6035