

Simulation model to the online course advising

Alejandro Felipe Zárate Pérez¹ and Idalia Flores de la Mota¹

¹ Universidad Nacional Autónoma de México, Av. Universidad 3004, Col Copilco Universidad, Coyoacán, Ciudad de México, 04510, México
alejandro.zarate@unam.mx

Abstract. The processes involved in online teaching have been inherited from traditional face-to-face courses, in almost every case we transfer the face-to-face to the virtual without making an analysis about how the things work in virtuality. In this paper we propose a simulation model for the course taught process, and from the model we will review a point that we assume to be definitive: the quantity of hours that an advisor dedicates to grading a course. Even the size of the student's group assigned to an advisor is defined by administrative policy but not necessarily by the needs of the training program, which one can be different according to the project's aim. In this paper we are going to model the course online process to the case of study, with this model will identify if the workload advisor is optimal or not.

Keywords: simulation, optimization, teaching analytics, e-learning analytics.

1 Introduction

In the last thirty years, technology has been advancing faster and this has also been reflected in education, an example of this is the modality called "Distance Education", which has gained more strength, as it is more accessible thanks to innovative technologies. This modality has the purpose to give education access to diverse sectors that have not been able to be attended, due to situations such as geographic, employee, time, among others.

Nowadays, with the incorporation of ICT (Information and Communication Technologies), it is possible to glimpse the scope that these represent for distance education, thus playing an essential role, because of the application of these recent technologies to the educational and training field, what is called "e-learning".

E-learning is a way of using ICT as a means of distribution for educational materials and other services, in which there is also an interrelation between teachers and students. Thus, in this new teaching-learning environment, web technology is used through the Internet.

Within education we find two types of education: face to face and continuing education. In this paper we will focus only on the second: continuing education.

There are a lot of definitions about continuing education, some of which vary according to the country to which we refer. However, for this paper we will take the UNAM definition of continuing education:

“It is an educational modality designed, organized, systematized, and programmed that complements the curricular formation and deepens and broadens knowledge in all fields of knowledge; it trains and updates professionally and is aimed at the university community and the public.”

The Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC) part of Universidad Nacional Autónoma de México, offers online continuing education courses, mainly in computing, through the Coordination of Continuing Distance Training, which are aimed to the public, the university community and institutions and companies that request them.

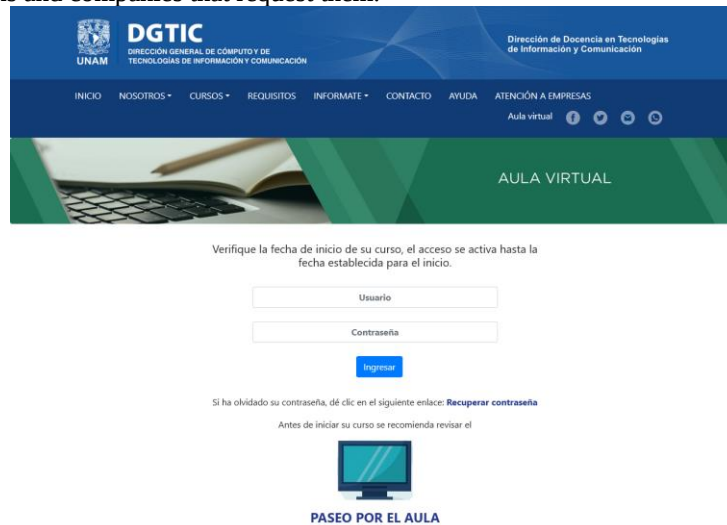


Fig. 1. Online course website

2 State of art

An e-learning advantage is that it will save time and money (Beetham H & Sharpe R, 2010), but we can find case where the e-learning is consuming more time and expensive than traditional teaching. (Laurillard, 2007) This situation applies not only the students, also the teacher needs more time to grade, solve questions, messages to the students.

Although the traditional (face to face) an online teaching method are valuables, they have some differences, meanwhile the traditional process of learning that is based on joint work with all students at the same time and where group work or individual training is used only as an additional form of training, the online training's dominant is independent learning and independent work. (Signe, Dace, & Edgars, 2015)

In the case we are studying, for each of the courses taught, the course materials and activities are developed according to a didactic planning made for the needs of each course. This implies that all groups in the same course see the same materials and perform the same activities during the course. (Zarate Perez & Flores de la Mota, 2023)



Fig. 2. Example of online course materials.

By using a model of this type, the online courses that are taught do not require the figure of a teacher who oversees transmitting the knowledge, since this is planned through the materials and activities that are developed. For this reason, for each group that is opened, an Advisor is assigned, who is only in charge of resolving doubts and evaluating the necessary activities. (Zarate Perez & Flores De la Mota, Assignment model for courses online, 2022)

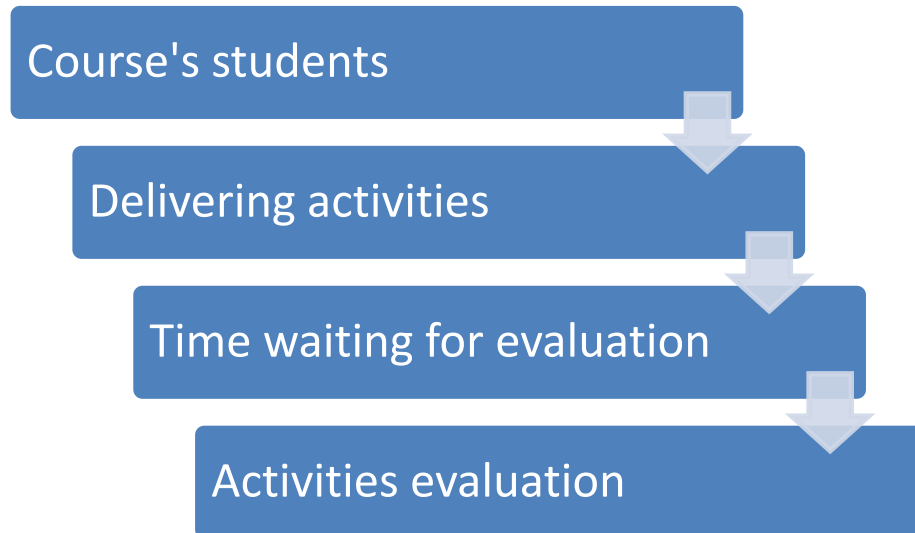
It is also necessary to mention that the activities of these courses are planned on a weekly basis, i.e., the materials and activities must meet the learning aims each week.

Although in some cases this advisor may take part in the development of the materials and activities, in general, the advisors assigned to the courses are not involved in the development of the course materials and activities and are limited only to the resolution of doubts and evaluation of the activities.



Fig. 3. Topic evaluation example

Under this advisors' model we can define the process as follows:



Each of these parts are described as follows:

Course's students. We refer to students who actively perform the activities showed before submitting an activity that requires evaluation. Such activities are materials consultation, forums participation, etc.

- **Delivering activities.** This part corresponds exclusively to the delivery by the student of the activity that will be evaluated for the corresponding week. As mentioned before, the activities are planned by week, so the submitting deadline for the evaluation will be at the end of the week, but it does not necessarily happen this way, since students submit activities from the beginning of the week until the end of the week.
- **Time waiting for evaluation.** In this waiting time we refer to the time elapsed between a student's submission and the time when an advisor starts grading; this is because students can submit activities at any time but must wait for the advisor to log in to the course site to grade. In this case, since advisors log in daily, the maximum time for an advisor to begin grading an activity is 24 hours.
- **Activities evaluation.** Here, as its name shows, the advisor evaluates the activities delivered by the student, the time it takes for an assessor to evaluate an activity depends on the delivery made by the student.

3 Modeling problem

At a technical level, we can affirm that all the systems necessary to teach online courses use a database for their correct operation, in which the information of the lessons, activities and even the participants' grades are stored.

However, not only does it store didactic information, but it also collects information on the interactions of the participants within the platform, this information can be useful because through it we can find how the teaching-learning process is developing.

The operation of online courses is done through a Moodle LMS, the entire course is conducted within this platform, so all activities are recorded in the system database, among these activities are review of materials, delivery of assignments, participation in forums, etc.

From this information we can make observations of each of the parts of the process mentioned above, with which it is possible to statistically model the behavior of each of these stages.

Therefore, in general, the simulation model will be as follows:

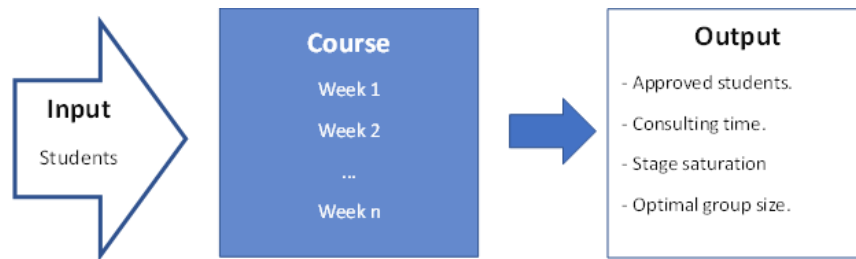


Fig. 4. Simulation model

This implies that the simulation model takes a certain number of students, who enter the course, perform their weekly activities and at the end we could answer questions such as: How many students pass? how long was the counseling time? in which phase is the model most saturated? what is the optimum group size?

However, as we can see within the course modeling students must pass the activities for each week, within each week the simulation model is as follows:

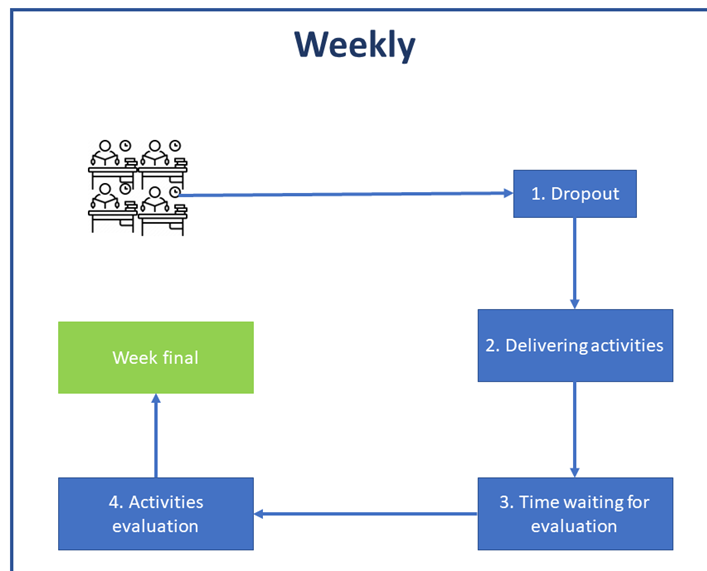


Fig. 5. Weekly model

Therefore, the simulation model of the course is composed of the individual simulations of each week until the end of the course.

4 Example

To exemplify this model, we are going to calculate the counseling time employed, this is important because up to now each advisor has been given a time of 20 hours. This time is decided by the policies of the institution, but it has not been proven that this time is sufficient.

The data that we will use in each phase are described below. These data were taken during a training program given to an institution where 4-week courses were given with an assigned counseling time of 20 hours for each course.

1. **Dropout (D).** The observed dropout rate was 5%, which is extremely low compared to the rates seen in courses open to the public.
2. **Delivering activities.** For this phase, we will use the following frequency table, which shows how the activities were delivered according to the day of the week.

Table 1. Frequency table by day

Day	Frecuency
Monday	1.02%
Tuesday	3.06%
Wednesday	7.33%
Thursday	13.92%
Friday	26.03%
Saturday	48.59%
Sunday	0.05%

3. **Waiting time for evaluation (TE).** According to the observed data we will use an exponential distribution with $\lambda = 12$. This occurs because students make the activity submissions at the end of each day and the assigned advisor generally reviews these submissions the next morning.
4. **Grading (TG).** As mentioned before, this phase refers to the time it takes for an assessor to rate an activity, for this part we will use a normal distribution with $\mu = 25$ and $\sigma = 15$.

We also used groups of 20 students, so the result of a simulation is shown below.:

Table 2. Simulations results per week

W eek		Mon- day	Tu esday	Wen sday	Thu rsday	Fri day	Sat- urday	Su nday
1	Stu- dents	1	1	0	5	5	8	0
	TE	7	19	0	8	16	11	0

	TG	36	4	0	78	11	153	0
						4		
2	Stu- dents	0	0	3	4	3	9	1
	TE	0	0	20	6	8	8	15
	TG	0	0	94	170	89	155	69
3	Stu- dents	0	0	1	0	5	13	0
	TE	0	0	4	0	25	13	0
	TG	0	0	16	0	10	351	0
						3		
4	Stu- dents	0	0	2	1	5	11	0
	TE	0	0	4	11	7	15	0
	TG	0	0	84	29	16	254	0
						5		

In summary, we have the following:

Table 3. Simulation results for course

	Week 1	Week 2	Week 3	Week 4
Final students	20	20	19	19
TE' average	8.79	8.05	5.98	5.19
TG's sum	383.98	577.28	470.42	532.05

This shows that at the end of the course **19 students finished** the course, the average **waiting time** for grading was **7 hours** and the time used by the **advisor to grade** the activities was **32.73 hours**.

We run this simulation process 20 times, the results are presented in the following table.

Table 4. Simulations results

Simula- tion	Final students	TE	TG
1	19	7.00	32.73
2	18	9.83	33.91
3	18	7.83	34.97
4	18	7.55	33.01

Simula- tion	Final students	TE	TG
5	19	7.02	35.14
6	19	5.85	37.70
7	18	7.35	31.96
8	17	6.82	31.93
9	16	6.71	25.06
10	19	7.50	30.57
11	17	8.84	30.39
12	19	7.74	30.38
13	19	8.69	29.11
14	18	7.86	31.57
15	19	8.02	31.01
16	19	6.73	31.17
17	15	6.59	26.99
18	19	7.32	31.45
19	18	7.04	27.74
20	19	8.28	31.96

The results of these simulations show that the average waiting time was 7.53 hours and the average time to grade was 31.44 hours. In this case we can see that the minimum time to grade was 25 hours when the student dropout rate was the highest.

However, the 20 runs were only to exemplify the results because they are no representative. So, the question here will be, how many runs we need to consider stable the simulations results?

To answer this question, we had to run 10 simulations, then 20, 30, up to 1,000. In all simulations set we calculate the average final students, waiting time and grading time. With this data we obtain the next table:

Table 5. Simulations results

Simulations	Average Final students	Average TE	Average TG
10	16.1	7.07	28.87
20	16.4	7.28	28.32
30	16.83	7.42	30.55
40	16.63	7.28	29.57
50	16.2	7.31	29.41
...	..	..	..
1000	16.25	7.3	29.36

The next step is normalizing the data, to do it we obtain the difference between the simulations n less the simulations $n-1$ and then calculate the absolute value. After this process we have 999 rows.

Finally, we calculate the moving average of 5 with the differences, we do this to smooth the data and then the graph will look a clearly tendency.

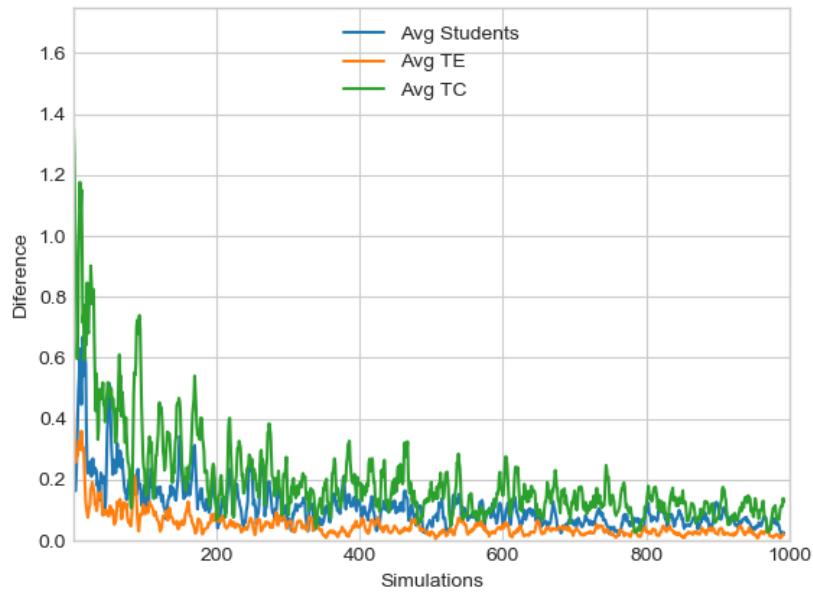


Fig. 6. Resulting graph from 1 to 1,000 simulations.

We observe that over 800 simulations the differences in the 3 cases maintain under 0.2, this is small value, and we will consider acceptable. Therefore, to the next scenarios analysis we will use exactly 800 simulations to calculate the values.

The results after **800 simulations** are **16 final students**, **7.32 hours to average waiting time** and **29.39 hours to average grading time**. We have to remember the initial parameter: 20 students at the beginning, a Poisson distribution with lambda 12 for the waiting time for evaluation and a Normal distribution with media 25 and sigma 15 for the grading time; also, we assume that an advisor have assigned 20 hours for grade students.

But the obtained results shows that the advisor needs 29 hours for grade, i.e., 50% more time that assigned originally. What should we do to reach 20 hours or less?

4.1 Scenarios

We are going to find the optimal values without exceeding the 20 hours for grading time in the course. For this case we have two options: decrease the size group or decrease the grading time per activity.

Size group

In this case we are going to move the size group from 10 to 20 and analyze its behavior with the grading time hours. The results are shown in the next figure.

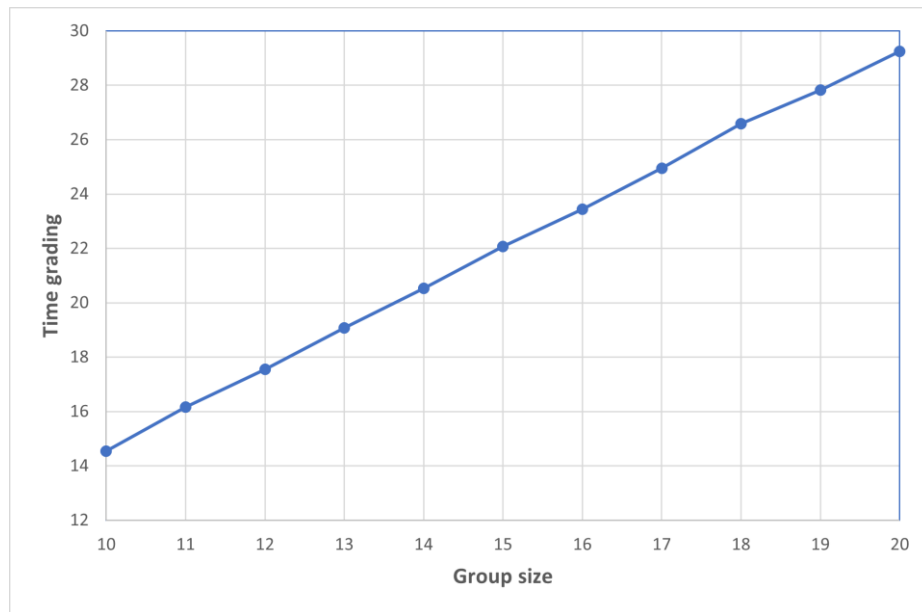


Fig. 7. Total grading time by size group

According with these results we can observe that the optimal group size is between 13 and 14 students, the total grading time between 19.07 and 20.52 hours. Therefore, if the group size is reduced between 30% and 35% its possible do not exceed the time grading assigned.

Grading time per activity.

For this scenario we will decrease the grading time, we will decrease the parameters for the Normal distribution, the values used, and the results are in the next table.

Table 6. Results moving the Normal distribution parameters.

Media	Sigma	Grading time
25	15	29.34
20	10	23.41

18	13	21.13
17	12	20.00
16	11	18.89
15	10	17.54
10	5	11.76

Therefore, the optimal value is between Normal(16, 11) and Normal(17,12) with grading time between 18.89 and 20 hours.

However, this case is the most difficult because the change in the Grading time per activity implies that its necessary a change in the activity itself, and possibly will be necessary a pedagogical review in every activity, which it's a task that requires time, analysis, and a test set to verify the pedagogical validity in every activity.

5 Conclusions

The results obtained during the simulations have allowed us to find the hours that an advisor needs to the grade students' activities, which differs from the official hours allocated by an additional 50%. This is relevant because the assessor is exceeding by far the working time, which must be readjusted according to the resulting data.

In this case, more advisor time could be given, or the student group size could be reduced. For the latter case, the simulation model can be used in such a way that instead of fixing the number of students, what would be fixed would be the maximum number of hours to run the necessary simulations to find the student optimal group necessary to use only 20 hours of advisor time.

However, this is not the only case where this simulation model can be used, as it can be used to set up before the start of a training project the necessary resources to complete without problems: advisor's quantity, time's advisor, group size, among others.

Likewise, the model described in this work can be used during the course or training program, with the purpose of predicting upcoming events by feeding the model with those generated in real time, i.e., using historical and current data to identify behaviors and make the necessary decisions if they will prevent the achievement of the established objectives.

One of the disadvantages of the model lies in the observed data, because it is necessary to make an adequate collection of them, to make the corresponding processing with them and to find the probabilistic models that best adapt to them. This is so that the simulation results are as close as possible to the real situation.

On the other hand, the greatest advantage of the model is the flexibility it gives us to answer questions such as: What is the optimal group size? How much time is dedicated to grading? How many advisors do we need for N students? among others. These questions can be solved using exact methods (queuing theory, dynamic programming), but these need assumptions that cannot be changed quickly, and it would be necessary to develop a model for each specific situation or training program.

Finally, the model allows us to answer questions quickly with only minor adjustments to the variables and to reuse it in different courses and/or training programs that behave differently.

References

1. Ndukwe, I.G., Daniel, B.K. Teaching analytics, value and tools for teacher data literacy: a systematic and tripartite approach. *Int J Educ Technol High Educ* 17, 22 (2020). <https://doi.org/10.1186/s41239-020-00201-6>
2. S. Robinson, "A tutorial on simulation conceptual modeling," 2017 Winter Simulation Conference (WSC), Las Vegas, NV, USA, 2017, pp. 565-579, doi: 10.1109/WSC.2017.8247815.
3. C. J. Lynch, R. Gore, A. J. Collins, T. S. Cotter, G. Grigoryan and J. F. Leathrum, "Increased Need for Data Analytics Education in Support of Verification and Validation," *2021 Winter Simulation Conference (WSC)*, Phoenix, AZ, USA, 2021, pp. 1-12, doi: 10.1109/WSC52266.2021.9715485.
4. Michael G.H. Bell, Stochastic user equilibrium assignment in networks with queues, *Transportation Research Part B: Methodological*, Volume 29, Issue 2, 1995, Pages 125-137, ISSN 0191-2615, [https://doi.org/10.1016/0191-2615\(94\)00030-4](https://doi.org/10.1016/0191-2615(94)00030-4).
5. Ricardo Cao Abad, *Introducción a la Simulación y a la Teoría de Colas*, Colección Netbiblio, 2002, ISBN 8497450175
6. Soon, W. M., & Ang, K. C. (2015). Introducing queuing theory through simulations. *The Electronic Journal of Mathematics and Technology*, 9(2), 152-165.
7. Fuller, Daniel & Arruda, Edilson & Martins Ferreira Filho, Virgilio Jose. (2019). Learning-agent-based simulation for queue network systems. *Journal of the Operational Research Society*. 10.1080/01605682.2019.1633232.
8. Zárate Pérez A.F., and Flores de la Mota I. (2022)., Assignment model for courses online. *Proceedings of the 34th European Modeling & Simulation Symposium (EMSS 2022)*. , 038 . DOI: <https://doi.org/10.46354/i3m.2022.emss.038>
9. Pérez, A. F. Z., & de La Mota, I. F. (2023). Un modelo para el seguimiento de cursos de capacitación. *Brazilian Journal of Development*, 9(2), 8464–8480. <https://doi.org/10.34117/bjdv9n2-146>
10. Signe, B., Dace, B., & Edgars, S. (2015). The Model for Balancing Learning Workload. *Procedia Computer Science*, 113-118.
11. Laurillard D. 2007. Modelling benefits-oriented costs for technology enhanced learning. *Higher Educ* 54:21–39.