

MODELACIÓN ESTADÍSTICA PARA EL APRENDIZAJE DE LA CORRELACIÓN Y REGRESIÓN LINEAL

Elkin Osorio Amaya, Santiago Inzunza Cáceres y Silvia Evelyn Ward

Este trabajo muestra los resultados del uso de una actividad generadora de modelos para el aprendizaje de la correlación y regresión lineal. Con su resolución, los participantes adquirieron, ampliaron y refinaron sus conocimientos y habilidades estadísticas, organizando, analizando y representando datos, estableciendo correlaciones, determinando líneas de regresión y coeficientes de ajuste; asimismo, con los modelos planteados, concluyeron, explicaron y describieron el fenómeno. Además de establecer la actividad como herramienta pedagógica y elemento útil para la investigación en entornos virtuales, el estudio muestra carencias en la modelación de los estudiantes, vinculadas a la planeación previa al abordaje de la problemática.

Términos clave: Actividades generadoras de modelos; Correlación lineal; Estadística educativa; Modelación estadística; Regresión lineal

Statistical modeling for learning correlation and linear regression

The research shows the results of using a model-eliciting activities for learning correlation and linear regression. With its resolution, the participants acquired, expanded, and refined their statistical knowledge and skills, organizing, analyzing, and representing data, establishing correlations, and determining regression lines and coefficients of determination. Also, with the proposed models, they concluded, explained, and described the phenomenon. In addition to establishing the activity as a pedagogical tool and a useful element for research in virtual environments, the study shows deficiencies in the modeling of students linked to planning prior to addressing the problem.

Keywords: Educational statistics; Linear correlation; Linear regression; Model-Eliciting Activities; Statistical modeling

Modelagem estatística para aprendizagem de correlação e regressão

Este documento mostra os resultados da utilização de uma actividade geradora de modelos para aprendizagem da correlação e regressão linear. Ao resolvê-lo, os participantes adquiriram, expandiram e aperfeiçoaram os seus conhecimentos e competências estatísticas, organizando, analisando e representando dados, estabelecendo correlações, determinando linhas de regressão e coeficientes de ajustamento; também, com os modelos propostos, concluíram, explicaram e descreveram o fenómeno. Para além de estabelecer a actividade como ferramenta pedagógica e elemento útil para a investigação em ambientes virtuais, o estudo mostra deficiências na modelação dos estudantes, ligadas à planificação antes de abordar o problema.

Palavras-chave: Actividades de construção de modelos; Correlação linear; Estatística educacional; Modelação estatística; Regressão linear

La correlación y regresión son conceptos fundamentales para el estudio de fenómenos en los que interesa analizar la relación entre dos o más variables (Engel y Sedlmeier, 2011). El razonamiento correlacional, la covariación y las relaciones entre datos multivariados son la base para el procesamiento de información que el individuo requiere en situaciones de incertidumbre (Gea, 2014). Por tanto, surge la necesidad de analizar datos provenientes de varias variables, situaciones que están presentes a lo largo de la vida del ser humano, ya sea en situaciones cotidianas o profesionales. Por ejemplo, un individuo realiza inconscientemente correlaciones entre la cantidad de energía que le proporciona cierta bebida o comida y el tiempo requerido para resolver alguna actividad, considerándolas como base para decidir qué consumir en función de la actividad que realizará; de manera similar, se puede correlacionar el clima con el consumo de energía eléctrica o el peso de un paquete con el costo del envío, por mencionar algunos otros casos.

Del análisis más profundo de la relación entre variables se desarrolla la idea de regresión, como método estadístico para describir, explicar y predecir el comportamiento de la relación entre variables (Seber y Lee, 2003). Cuando interesa ir más allá de determinar la existencia de la correlación, se pueden prever posibles resultados y verificar si los datos tienen alguna tendencia, favoreciéndose de la información presentada o vislumbrando incoherencias entre el discurso y los datos presentados. Estos aspectos evidencian la importancia de la comprensión de las personas sobre estos temas, particularmente, donde se analicen situaciones problemáticas reales y significativas para el individuo.

En la estadística, la modelación es un campo de investigación que se centra en la enseñanza y aprendizaje del proceso de razonamiento estadístico de los estudiantes, dentro de los paradigmas de la perspectiva de modelación matemática

(Pfannkuch et al., 2018). El foco es el desarrollo de procesos investigativos por parte de los estudiantes, abordando un fenómeno del mundo real con el fin de modelarlo, describirlo, explicarlo y, si es el caso, plantear soluciones a las problemáticas que presenta (Patel y Pfannkuch, 2018). Con este enfoque se potencia el desarrollo de los conocimientos de los estudiantes por medio de los ciclos investigativos, en los cuales se hace una aproximación al problema, se comprende, se traza un plan, se recopilan y analizan datos y se elaboran conclusiones al respecto de la problemática (Wild y Pfannkuch, 1999).

En esta investigación nos planteamos como objetivo general analizar el proceso de modelación estadística que desarrollan los estudiantes a través del uso de actividades generadores de modelos (AGM) en la búsqueda de caracterizar este tipo de actividades para el aprendizaje de la correlación y la regresión lineal, identificar las posibles dificultades en su instrucción y determinar el potencial impacto en el desarrollo del conocimiento de los estudiantes. En específico nos planteamos las siguientes preguntas.

- ◆ ¿Cómo se caracterizan las AGM orientadas al desarrollo del conocimiento de los estudiantes en las temáticas de correlación y regresión lineal?
- ◆ ¿Qué dificultades vinculadas a la implementación de las AGM en el proceso educativo de los estudiantes existen?
- ◆ ¿Cómo se ve afectado el conocimiento de los estudiantes, relacionado con las temáticas de correlación y regresión, con la intervención educativa basada en la implementación de las AGM?

REFERENTES TEÓRICOS

En las últimas tres décadas, la introducción de la modelación matemática ha sido una característica común en las reformas al currículo matemático a nivel internacional (Kaiser et al., 2006). Sin embargo, su enfoque puede variar en función de la instrucción, la metodología y la forma en la que se conceptualiza la modelación. En particular, en nuestro trabajo se adoptó la vertiente contextual de la modelación, en la cual se propone el aprendizaje y el desarrollo de habilidades de los estudiantes a través de la resolución de problemas planteados dentro de situaciones cuidadosamente estructuradas y relacionadas con su contexto, con el objetivo de generar modelos matemáticos que permitan describir, explicar y predecir el comportamiento de la situación (Blomhøj, 2009).

El centro del proceso de comprensión de conceptos matemáticos está en la construcción misma del modelo, “herramienta conceptual de un sistema matemático desarrollado de una situación específica del mundo real” (Abassian et al., 2020, p. 7). De manera similar, este enfoque es la esencia de la correlación y la regresión lineal, construcción derivada de una situación en donde el comportamiento de una variable se relaciona con el comportamiento de otra

mediante un modelo que ayuda en la descripción, explicación y predicción de dichos comportamientos (Gea et al., 2014).

Desde esta perspectiva, las herramientas principales de trabajo son las actividades generadoras de modelos (AGM), (MEA por sus siglas en inglés para model-eliciting activities). El diseño de estas actividades se basa en seis principios descritos por Lesh y Doerr (2003) que se detallan:

1. El principio de la realidad. La situación debe parecer significativa para los estudiantes y conectar con sus experiencias de formación.
2. El principio de construcción de modelos. La situación debe crear una necesidad en los estudiantes para desarrollar constructos matemáticos significativos.
3. El principio de la autoevaluación. La situación debe permitir a los estudiantes evaluar los modelos provocados.
4. El principio de la documentación. La situación y el contexto deben requerir que el estudiante exprese sus pensamientos mientras solucionan el problema.
5. El principio de generalización. Los resultados deben ser generalizables a otras situaciones similares.
6. El principio de la simplicidad. La situación problema debe presentarse tan simple como sea posible sin perder de vista la necesidad de construir un modelo significativo para dar una explicación (p. 43).



Figura 1. Comparativa modelación y problemas tradicionales en palabras
(Adaptado de Lesh y Doerr, 2003, p. 4)

A diferencia de los problemas tradicionales, donde la situación planteada se ofrece acompañada de fórmulas, datos puntuales y palabras clave que le indican al

estudiante cómo proceder y donde solo se requiere que este decodifique el enunciado; en la modelación se requiere que el estudiante describa matemáticamente una situación significativa (Figura 1). Estas actividades involucran compartir, manipular, modificar y reusar herramientas conceptuales para la construcción, descripción, explicación, manipulación, predicción o control de sistemas matemáticamente significativos (Lesh y Sriraman, 2010).

El proceso de construcción de estos modelos, cuando se orienta en elementos estadísticos, se realiza a través de ciclos investigativos (Wild y Pfannkuch, 1999). En tales ciclos, el estudiante se aproxima al problema planteando preguntas de investigación relacionadas con la situación del mundo real que enfrenta, elaborando con base en ellas un plan para resolverlas. Dichas preguntas se transforman en datos relevantes para ser recolectados y analizados por medio de diversas herramientas y representaciones con el fin de obtener inferencias acerca de la problemática abordada, a su vez, obtienen conclusiones con las cuales puedan explicar, describir y predecir el comportamiento de la situación (Patel y Pfannkuch, 2018). Debido a las siglas de cada etapa (Problema, Plan, Datos, Análisis y Conclusiones), este ciclo se denomina PPDAC (Figura 2) y es considerado como base para el análisis del proceso de modelación del estudiante. Sí, por medio del tránsito del estudiante por las distintas etapas de este ciclo es posible identificar el avance de su entendimiento sobre los conceptos trabajados, evidenciando su comprensión con la profundidad del análisis, el adecuado uso de herramientas, el nivel de refinamiento del modelo y la calidad de las conclusiones que obtiene.

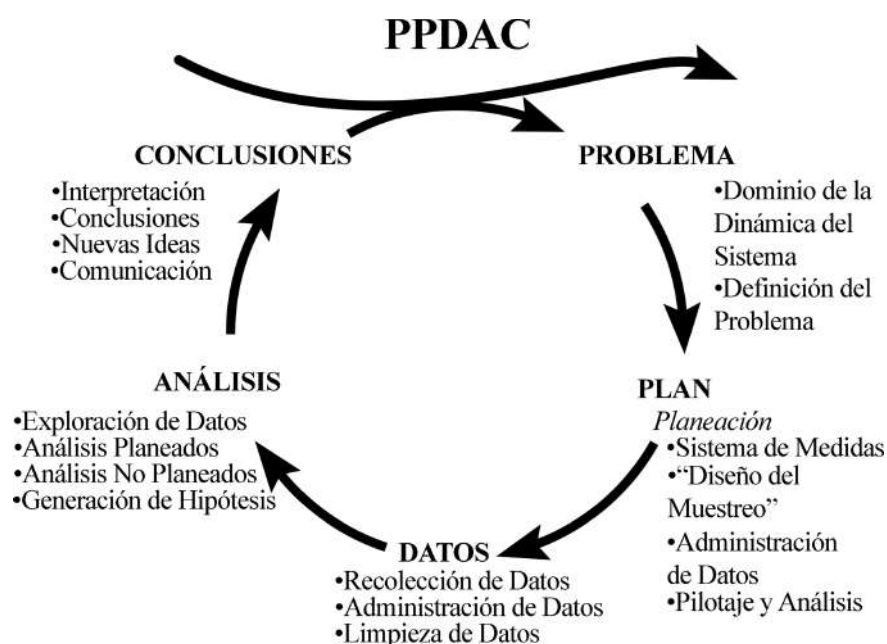


Figura 2. Ciclo investigativo PPDAC (Adaptado de Wild y Pfannkuch, 1999, p. 226)

De manera sintética, el estudiante, al transitar a través de un ciclo investigativo:

- ◆ Aborda el problema definiendo sus características, elementos involucrados y su estructura, hasta poseer un dominio de la dinámica del sistema.
- ◆ Realiza un plan en el cual se toman en cuenta aspectos como el sistema de medidas, la muestra para el estudio, el pilotaje, entre otros.
- ◆ Recolecta datos, organiza, clasifica y determina su relevancia.
- ◆ Analiza los datos, con base en su plan o en función de las eventualidades, utilizando diferentes representaciones que faciliten su propósito y generando así hipótesis relacionadas con la situación.
- ◆ Establece sus conclusiones, interpretaciones o el planteamiento de nuevas ideas que permitan abordar el fenómeno nuevamente.

ANTECEDENTES

La perspectiva de modelos y modelación en las últimas décadas ha tenido un desarrollo importante tanto en sus aspectos conceptuales como siendo una propia estrategia de enseñanza. Por ejemplo, la serie de libros de la editorial *Springer: Models and Modeling in Science Education* lleva 12 volúmenes desde 2005 a 2019. En su última edición, se aborda la modelación como una competencia fundamental a alcanzar para desarrollar conocimientos en las ciencias (Upmeier zu Belzen et al., 2019), al mismo tiempo, se ofrecen guías prácticas basadas en

evidencias para la evaluación y promoción de esta competencia (Constantinou et al., 2019).

Es en este sentido en el que se desarrolla esta investigación, en la búsqueda de evidencias que fungan como fundamento para guiar la implementación de actividades de modelación.

Las investigaciones de este tipo en el área estadística también proliferaron. Existen números de revistas prestigiosas enteramente dedicados a la modelación, como la revista *ZDM-Mathematics Education* que, en su número 50, presentó múltiples artículos sobre modelación estadística, mayormente implementando AGM con apoyo de software especializado. Un ejemplo de estos es el trabajo de Aridor y Ben-Zvi (2018), quienes describen cómo el pensamiento relacionado (*aggregate reasoning*), considerado núcleo del razonamiento estadístico, puede ser desarrollado en los estudiantes por medio de actividades de modelación de fenómenos reales. Particularmente, este trabajo estuvo enfocado en el pensamiento relacionado con muestras y muestreo, utilizando actividades que involucran análisis de datos e inferencia estadística por medio del software especializado TinkerPlots®. Estas actividades se basaron en los planteamientos de Lehrer y English (2018), quienes a su vez citan a Lesh y Doerr (2003), lo que sugiere que poseen fundamentos similares a las AGM. En comparación con nuestro trabajo, estos planteamientos se diferencian en la especificidad acerca de los ciclos de modelación por los que transita el estudiante al desarrollar la actividad. Para Aridor y Ben-Zvi (2018), debido al carácter probabilístico de los datos utilizados, el estudiante debe ir “pasando del mundo real al mundo probabilístico y de vuelta” (p. 4), en tres ciclos que se basan en muestreo de datos, inferencias y de vuelta a muestreo de datos. En nuestro caso, las actividades se plantean para que el estudiante recorra el ciclo de modelación estadística PPDAC, pues en este proceso de construcción del modelo es posible observar el avance segmentado por etapa, lo que permite establecer fallas puntuales de instrucción, planeación, procesamiento, ejecución o conclusión, cuando se trabaja con las AGM.

Más recientemente, Aridor y Ben-Zvi (2019) abordan la covariación para el desarrollo de este pensamiento relacionado, considerado por ellos como una “visión global de los datos considerando características agregadas de los conjuntos de datos y sus propensiones” (p. 74). De igual manera que la descrita anteriormente, se basan en actividades de modelación de fenómenos reales; con ello, se busca que el estudiante en su tránsito por los ciclos de modelación, haga inferencia estadística informal, la cual involucra los procesos de organización, clasificación, análisis, interpretación de datos y predicción con base en éstos. En este trabajo, los autores utilizaron una actividad basada en aviones de papel en la cual los estudiantes construyen distintos de estos, con diferentes longitudes de alas y recolectan información sobre el tiempo de vuelo y distancia recorrida, lo que los lleva a plantear modelos que describen el comportamiento con base en una regresión lineal de una nube de puntos que representa las variables relacionadas con el avión de papel. A pesar de coincidir con nuestro trabajo en la temática, el

tipo de herramienta y la forma de análisis, Aridor y Ben-Zvi (2019) se alejan de nuestro objetivo y la forma de llegar a él; los autores con este estudio de caso concluyen que pueden iniciar a los estudiantes “con una visión local de la covariación” y establecer un “entorno multifacético” (p. 92) de aprendizaje para desarrollar este pensamiento relacionado; esto les daría las bases para desarrollar su pensamiento estadístico. No obstante, en nuestra investigación buscamos desarrollar el conocimiento necesario en los estudiantes para la generalización de estos análisis covariacionales de correlación y regresión lineal con base en el modelo generado y generalizado a partir del desarrollo de la AGM.

Dentro de las publicaciones de la revista *ZDM-Mathematics Education* se encuentra el marco de modelación estadística planteado por Patel y Pfannkuch (2018), definido como SWAMTU (por sus siglas en inglés: *situation, want to know, assumptions, model, test y use*) y utilizado para caracterizar el razonamiento de los estudiantes. Los investigadores implementaron AGM y secuencias instruccionales *follow-up task* (las que se basan en tareas anteriores llevadas a un nivel más alto de complejidad), con la finalidad de que los estudiantes, mediante el tránsito iterativo por el ciclo de modelación planteado, desarrollaran su razonamiento estadístico. Estos autores utilizaron TinkerPlots® para que sus estudiantes procesaran los datos y, aunque determinaron que su modelo es viable para describir el desarrollo del razonamiento en los estudiantes con las AGM, también señalaron que es necesario realizar análisis comparativos con diferentes marcos y diferentes enfoques y de esta manera capturar todo el espectro de pensamientos y procesos que desarrollan los estudiantes.

Estas cuestiones son planteadas detalladamente por Pfannkuch et al. (2018) en su análisis sobre la modelación estadística. También discuten sobre las AGM como herramientas que abordan esta orientación a cabalidad y con excelentes resultados. Estos investigadores señalan dos cuestiones importantes en el proceso de implementar la modelación estadística: el rol del contexto y la innovación tecnológica. En relación con lo primero, los autores señalan que “sin contexto la modelación estadística no puede ocurrir” (Pfannkuch et al., 2018, p. 1115). Básicamente, hasta los datos en bruto pierden el sentido sin un contexto al que se relacionen; esto hace que recolectarlos, organizarlos, clasificarlos, representarlos y analizarlos no tenga significado alguno. No solo esto, pues si el contexto no resulta relevante para el estudiante, se compromete todo el proceso de aprendizaje. Sobre lo segundo, se describen los grandes cambios que produce la integración de tecnologías innovadoras en el proceso, por ejemplo, con la puede transformarse el énfasis en cómo se construyen gráficas a mano de un pequeño conjunto de datos, a centrarse en la exploración e indagación de relaciones en grandes conjuntos multivariados de datos. Con base en estas consideraciones se planteó el contexto de una de las AGM, la del calentamiento global, situación que puede considerarse relevante para cualquier población. Además, se pautó el procesamiento de los datos en el software de preferencia de los estudiantes, evitando cuestiones procedimentales y permitiendo así un enfoque en el proceso de modelación.

En síntesis, se observa que la perspectiva de modelación para el desarrollo de los conocimientos de los estudiantes tiene gran potencial mediante la implementación de las AGM. Las características de éstas propician escenarios para la relación de variables entorno a la descripción, explicación y predicción de fenómenos, el planteamiento de éstas en contextos significativos para el estudiante vuelve relevante su resolución y mejora el compromiso de estos.

METODOLOGÍA

La investigación aquí presentada se planteó como una ingeniería didáctica, enfocando el trabajo en la micro-ingeniería debido a lo particular de la temática, pues se estudia el resultado de la implementación única de la AGM en una intervención educativa (Artigue et al., 1995). La metodología para el análisis de los datos fue de tipo cualitativa, centrada en la identificación y comprensión del proceso de modelación estadística de los estudiantes participantes, por medio de la observación e interpretación de sus respuestas, productos generados en cada etapa de sus ciclos de modelación y el refinamiento y generalización de los modelos construidos. La recolección de los datos se hizo mediante los productos escritos, notas de observación y las grabaciones de las clases.

Participantes

En el estudio participaron nueve docentes de matemáticas en formación de Colombia. Al finalizar su pregrado, su educación los prepara para ser profesores a nivel de primaria, secundaria y bachillerato. Estos estudiantes universitarios estaban en el quinto semestre de su carrera (la cual tiene nueve semestres) y tenían entre 20 y 23 años. En este punto de sus estudios ya habían cursado la única asignatura de Estadística considerada en el currículum de la licenciatura. Sin embargo, solo uno de ellos afirmó tener conocimiento al respecto de la correlación y regresión lineal, proviniendo éste de sus estudios previos en bachillerato.

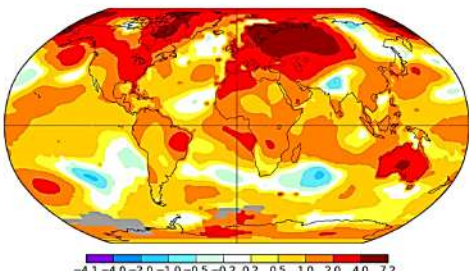
¿El final está cerca? Esto nos dicen los científicos y sus datos.

Redacción InforMex 6 febrero, 2020 Sin categoría

¿Cuántos años tienes? ¿15? ¿20? ¿30? ¿más? Pues déjame decirte algo que te impresionará. Si naciste en 1990, ¡la temperatura desde ese entonces a el día de hoy prácticamente se ha duplicado! Si, si, como lo lees.

Según los datos de la [National Aeronautics and Space Administration](#) (NASA), si la agencia que tanto escuchas y ves en los programas espaciales, a temperatura promedio global (TPG) pasó de 45 grados celcius a 85, desde 1990 a 2018. **¿Desde que naciste cuánto ha cambiado la TPG?**

December 2019 L-OTI(°C) Anomaly vs 1951-1980 1.12



Fuente: Goddard Institute for Space Studies (GISS) Surface Temperature Analysis



origen antropogenico y dicen que el problema está en la radiación solar. Según los datos de [LASP Interactive Solar Irradiance Data](#) (LISIRD), el cambio de 1999 al 2000 en la radiación solar, fue un aumento de 1361.5023 puntos a 1365.4337. Esto también podría explicar el aumento de la temperatura global del planeta.

Esto ya empieza a sonar muy raro, hay mucha información, muchas fuentes, opiniones contradictorias y no se sabe qué pensar.

Si tuviéramos un grupo que se encargara de las cuentas, se podrían aprovechar los datos, pero al no tenerlo, sólo nos queda una opción: ¿y si nos ayudas? Si, tú que estás leyendo esto, animate y ayudáanos a entender si es que en realidad hay alguna forma de explicar este fenómeno.

Eso sí, considera que en este portal trabajamos con noticias, así que sé lo más claro posible, con gráficas, dibujos y colores si es necesario. Es más, si nos ayudas, publicaremos tu trabajo con un espacio en destacados. ¿lo harás?

Pues dejaremos toda la información disponible y las fuentes en un documento para su descarga aquí:

Datos NASA NOAA y LISIRD – INFORMEX [Descarga](#)

Muchas gracias por leernos y sigue informado con InforMex

Figura 3. AGM calentamiento global presentada al alumnado desde dispositivos móviles

Descripción de la AGM: calentamiento global

La AGM se estructuró con imágenes, textos, citas bibliográficas y datos reales de fuentes oficiales, como la Administración Nacional de Aeronáutica y el Espacio (NASA), la Oficina Nacional de Administración Oceánica y Atmosférica (NOAA) y *LASP Interactive Solar Irradiance Data Center* (LISIRD). Esta actividad es presentada a los estudiantes a través de internet, en un blog de noticias ficticio (<https://informex.data.blog/2020/02/06/agm-1/>), al cual puede accederse desde una computadora o teléfono móvil (variando únicamente su tamaño). En la figura 3 muestra la versión móvil de la AGM.

Esta AGM se construyó teniendo en cuenta cada uno de los principios de diseño (Doerr y English, 2003). Estos principios permiten asegurar que las situaciones propuestas simulan problemas de la vida real, implementan de manera útil el razonamiento matemático y delimitan la naturaleza de los conocimientos y habilidades que son necesarios para abordar satisfactoriamente la situación (Lesh et al., 2000).

- ◆ Principio de la realidad. La AGM calentamiento global se basa en datos reales y las fuentes utilizadas son confiables, además, como son organizaciones que se dedican al estudio de estas variables, se mantienen actualizadas y pueden usarse posteriormente en alguna nueva versión de la actividad. El fenómeno descrito es una temática actual, vigente y que afecta a toda la sociedad en general. Sobre ella se habla en noticieros, redes sociales, políticas públicas, espacios académicos, entre otros.
- ◆ Principio de construcción de modelos. Esta AGM permite la construcción de un modelo lineal que describa el comportamiento entre las distintas variables, tales como año versus la temperatura promedio global (TPG), y cada una de las emisiones; o la TPG y cada una de las emisiones. Junto a esto, también es posible observar que no existe correlación entre la radiación solar y el comportamiento de ninguna de las otras variables (emisiones, años y TPG).
- ◆ Principio de la autoevaluación. Existe vasta información relacionada con el fenómeno, así que la selección de la variable que puede explicar mejor el comportamiento puede ser autoevaluada. Es posible encontrar artículos que apoyen sus posturas, así como otros que las contravengan. Así mismo, la pregunta en el título de la presentación de la AGM, “¿El final está cerca?”, resulta sugerente como método para evaluar la capacidad de predicción de los modelos planteados.
- ◆ Principio de la documentación. Se requiere que el participante haga el procesamiento pertinente de los datos, explique detalladamente sus planteamientos, use gráficos y demás. Se puede decir que la actividad permite que los sujetos externalicen sus pensamientos, conocimientos, procesos, refinamientos y conclusiones en formas y momentos que pueden ser recopilados para su posterior análisis.
- ◆ Principio de generalización. Una vez definida la forma de representación apropiada de los datos de las variables que se desean relacionar, los estudiantes pueden ir probando con las demás variables y ver, que, aunque los modelos lineales identificados difieren en sus constantes, presentan la misma estructura como ecuación de una línea recta.
- ◆ Principio de simplicidad. La situación se plantea de manera simple, se presenta la existencia de un fenómeno y datos relacionados, pero distintas posibles explicaciones, haciendo falta de un análisis estadístico que ayude a entenderlo y con el cual tomar una postura informada.

Implementación de la ACG

La puesta en marcha y análisis de la AGM denominada “calentamiento global”, se elaboró para ser implementada mediante un diseño instruccional básico de seis fases descrito por Lesh y Doerr (2003). Esta estructura se basa en secuenciar las interacciones del participante con la situación o fenómeno planteado. Primero se hace una presentación donde el participante interactúa con la situación y determina su relevancia y establece un grado de compromiso. Luego, la aborda individualmente y se aproxima a ella con sus supuestos. Después la enfrenta en compañía de los miembros de un grupo, haciendo que en la interacción intervenga su conocimiento articulado con el de ellos. Posteriormente consolidan sus productos y lo presentan al grupo curso, ampliando aún más la interacción con la situación. Finalmente, convergen en la formalización de la estrategia que se institucionaliza con apoyo del docente. Este proceso genera múltiples interpretaciones del fenómeno, las cuales los participantes pueden externalizar y evidenciar en su proceso cognitivo. A continuación se describe cómo se implementaron estas fases en este estudio.

- ◆ **Presentación.** Se realiza a través del blog de noticias InforMex (<https://informex.data.blog>), creado en la plataforma WordPress con el propósito de simular un portal noticioso donde la situación o fenómeno de la AGM es una noticia (Figura 4). Se les invita a los estudiantes a revisar dicho portal y ver la noticia; aunque, en caso de tener dificultades de acceso, se prepara la versión impresa de la AGM. Presentar de esta manera la AGM tiene como objetivo generar el sentido de realidad necesario para la actividad, además de promover el uso de herramientas tecnológicas, pues el portal es responsivo al uso de móviles, tabletas o computadoras.
- ◆ **Trabajo individual.** Dentro de la AGM se preparan preguntas que buscan captar la atención del estudiante y comprometerlos al desarrollo de la actividad. Se les pide a los participantes que trabajen en resolver los cuestionamientos e ir planteando los primeros acercamientos a la explicación del fenómeno. Dentro de la actividad se solicita la documentación de estas respuestas, trabajo que es capturado por el profesor titular e investigador para el posterior análisis del proceso del estudiante.
- ◆ **Trabajo en pequeños grupos (tres estudiantes).** Lo siguiente es reunirse en grupos de tres estudiantes para compartir sus ideas y primeros planteamientos con el fin de desarrollar a cabalidad la AGM. El proceder de los participantes es documentado y recolectado para el análisis. En esta fase se observa la aparición de diferentes enfoques, procedimientos, resultados y modelos de los diversos grupos.
- ◆ **Consolidación y presentación.** Aunque se trabaje en pequeños equipos, el grupo en general suele brindarse apoyo entre ellos, esto enriquece los resultados y facilita la construcción de los modelos. Una vez se planteen las soluciones de los grupos, se determinan cuáles de estos, en función de la

divergencia de procedimientos y resultados, pasan a explicar su solución ante el grupo en general. Esto reafirma sus procedimientos y favorece a la refinación de sus conclusiones.

- ♦ Institucionalización. Con base en los elementos que presentaron los estudiantes (tablas, gráficos, fórmulas, modelos, etc.), se explica y fundamenta el sentido que tienen sus procedimientos, haciendo explícitos los conceptos, ideas, relaciones y métodos inmersos en el trabajo de modelación que realizaron.
- ♦ Informe final. Se solicita a los participantes el envío de propuestas con las cuales se pueda describir, explicar y predecir el fenómeno en cuestión. Se establecen pautas para la elaboración de este informe, principalmente de estilo e inclusión de elementos de apoyo como gráficas, ecuaciones, definiciones y demás aspectos que consideren relevantes, y luego se establece un límite de tiempo. Este informe final se trabaja de manera individual y permite observar una evolución de sus planteamientos iniciales.

Debido a la pandemia de Coronavirus ocasionada por el SARS-COV2, la implementación de las actividades fue afectada. Se pasó de las actividades en el aula de clases a las sesiones vía Google Meet, plataforma que se seleccionó entre las otras posibles (*Teams*, *Zoom*, *Skype*) por no requerir un proceso de instalación previo o algún registro de cuenta, además de tener acceso a sesiones gratuitas, programables y sin límite de tiempo. Lamentablemente, esto agregó un conjunto de limitaciones relacionadas con la comunicación, la recolección de información y el diseño instruccional.

En cuanto a la comunicación, los estudiantes cuando están en el aula de clases tienen cierto grado de privacidad al momento de expresar inquietudes al docente. Pueden hablar en tono bajo, llamar hasta el sitio donde está o ir donde el docente. Esto es un espacio más para la externalización de los procesos del estudiante, sus aproximaciones y refinamientos, los que influyen en su proceso de modelación y en su aprendizaje en general (Lesh y Doerr, 2003; Lehrer y English, 2018). En cambio, en una reunión virtual, estas oportunidades no están disponibles, todas las dudas que planteen los estudiantes serán escuchadas por sus compañeros, esto puede intimidar a los estudiantes, reduciendo así su participación (Moliní y Sánchez, 2019).



Figura 4. Presentación de la AGM en el portal InforMex

Como un derivado de esto, se tiene una reducción en la capacidad del investigador para la recolección de ciertos datos. Durante las actividades de modelación, el estudiante utiliza distintos medios para externalizar los subprocesos por los que transita; realiza bocetos de representaciones gráficas, operaciones en las márgenes, borradores de procedimientos, anotaciones sobre datos que consideran relevantes, entre otras cuestiones (Lesh y Doerr, 2003). En el aula, el docente puede acercarse al estudiante y acceder a esta información en el momento en que se realiza, cuestionarlo al respecto de ellas y obtener datos sobre su procedimiento; en cambio, en una sesión virtual, esto se limita considerablemente. Los estudiantes pueden decidir no mostrar al docente lo que hacen como boceto o anotaciones que no consideran presentables, y pueden alegar que no han hecho ninguna. Lo más cercano que se puede llegar a esto, es solicitar la presentación de su pantalla, en la cual se podrían observar cuentas en calculadoras, pruebas de generación de gráficas, las variables que seleccionan o los primeros párrafos que escriben.

Por último, en cuanto al diseño instruccional, se planeó trabajar en grupos pequeños de tres estudiantes como segundo momento del desarrollo de las actividades de modelación. Sin embargo, debido a las limitaciones técnicas de las plataformas para las reuniones virtuales, las mecánicas de trabajo colaborativo estuvieron limitadas a la creación de salas separadas donde los estudiantes,

después de desarrollar la primera sección de la actividad de manera individual, entraban a reunirse en las salas para cada grupo y, una vez terminada la sección de trabajo colaborativo, volvían con el grupo en general. Esto, además de generar incomodidades y confusiones técnicas, limitó la interacción entre los grupos de estudiantes, la cual es fundamental para el desarrollo de la actividad.

RESULTADOS

La actividad se presentó a los estudiantes como un artículo en el portal de noticias ficticio Informex. Se les proporcionó el enlace al que pudieron acceder inmediatamente, realizaron la lectura de la actividad y empezaron su trabajo individual. Dentro de la lectura se incluyeron un conjunto de preguntas para ubicar al estudiante en el fenómeno planteado: ¿Desde que naciste cuánto ha cambiado la temperatura promedio global?, ¿qué pensarías si te dijera que esto parece tener su causa en los humanos?, ¿produces tú alguno de estos gases? Estas preguntas fueron resueltas por los estudiantes y presentadas en un documento durante la sesión (en la figura 5 se pueden ver unos ejemplos de estas respuestas). De esta manera se puede observar que los estudiantes comprenden la situación presentada, la naturaleza del fenómeno expuesto y los elementos involucrados.

Carolina • ¿Cuántos años tienes?: 22 años. • ¿Desde qué naciste cuánto ha cambiado la TPG?: Ha variado 41 grados Celsius. • ¿Pues qué pensarías si te dijera que esto parece tener su causa en los humanos?: En parte es cierto por la contaminación que genera la industria creada por el hombre. • ¿Produce tú alguno?: Sí, dióxido de carbono.
Dairon • ¿Cuántos años tienes?: 19 años. • ¿Desde qué naciste cuánto ha cambiado la TPG?: 31. • ¿Pues qué pensarías si te dijera que esto parece tener su causa en los humanos?: Como seres vivos, somos los que más hacemos uso de los recursos de la naturaleza lo que genera un sobreconsumo de estos, lo cual implica un desgaste tanto del suelo como de la atmósfera. • ¿Produce tú alguno?: Dióxido de carbono.

Figura 5. Ejemplos de respuestas a las preguntas de ubicación

Entendimiento de la situación problema

El primer paso del ciclo investigativo (problema) hace referencia al abordaje y entendimiento de las características del fenómeno, los elementos involucrados, su estructura y dinámica (Lehrer y English, 2018). Con estas preguntas en la AGM, se focaliza al estudiante en el fenómeno y con la solicitud dispuesta (ayúdanos a

entender si es que en realidad hay alguna forma de explicar este fenómeno), se le centra en el objetivo, lo que permite que el abordaje al problema sea el buscado.

A continuación, los estudiantes cuestionaron sobre el cómo proceder, por ejemplo, la estudiante que llamaremos Maira como seudónimo, pregunta: “tengo una duda, o sea, ¿empezamos a trabajar de manera individual escribiendo sobre lo que leímos?”, lo cual se respondió diciéndoles que habría que ayudar a los miembros de la redacción del artículo, tal como lo solicitan. Esto puede leerse al final de la AGM, en la cual, sin establecerse rígidamente el proceso, se indica que la ayuda solicitada requiere explicaciones claras, con representaciones gráficas, dibujos y colores, tal como se muestra a continuación:

“[...] y si nos ayudas? Si, tú que estás leyendo esto, ámate y ayúdanos a entender si es que en realidad hay alguna forma de explicar este fenómeno. Eso sí, considera que en este portal trabajamos con noticias, así que sé lo más claro posible, con gráficas, dibujos y colores si es necesario”.

Los estudiantes no tuvieron experiencias anteriores con las AGM y cuando se les preguntó si habían resuelto una actividad con esta estructura, las compararon con actividades de resolución de problemas. Esto sugiere que esperaban un enunciado preciso sobre lo que se quería como solución. Para buscar evidencias sobre el plan que habían trazado, se les cuestionó sobre cómo resolverían la actividad, y se encontró que la única previsión que tomaron antes de proceder fue con respecto a las unidades de medida. Los estudiantes describieron que los datos sobre las emisiones de gases estaban en diferentes unidades (partes por millón, por billón y trillón) y que requerían normalizarlos en una misma escala para poder compararlos entre sí. Debido a esto, se les hizo saber que, en virtud de que las fuentes de los datos provienen de instituciones estadounidenses, los billones y trillones en los datos tenían connotaciones de miles de millones (109) y billones (1012) respectivamente.

Los estudiantes, en su mayoría, enfocaron sus esfuerzos directo en los datos sin detenerse en la planificación de su forma de trabajar. Los participantes solo observaron los datos y empezaron con el procesamiento de estos tomando como base los requerimientos de la actividad. Según el ciclo investigativo, quienes realizan el proceso de modelación estadística comenzarían con la organización, clasificación y determinación de la relevancia de los datos (Zieffler et al., 2018). Además, junto con la actividad se les proporcionó un conjunto de datos organizados por año, el cual tenía algunas posiciones vacías de manera intencional, años sin registros de datos; el objetivo de ello fue que intentaran completarlos, recopilando más información. No obstante, a pesar de notar esta falta de información, ninguno de ellos lo hizo, evidenciando una carencia de reflexión y planificación.

Primeras aproximaciones

Las primeras aproximaciones a los análisis de los datos se hicieron gráficamente. Los estudiantes utilizaron gráficos de barras, circulares, de línea y dispersión (Figura 6).

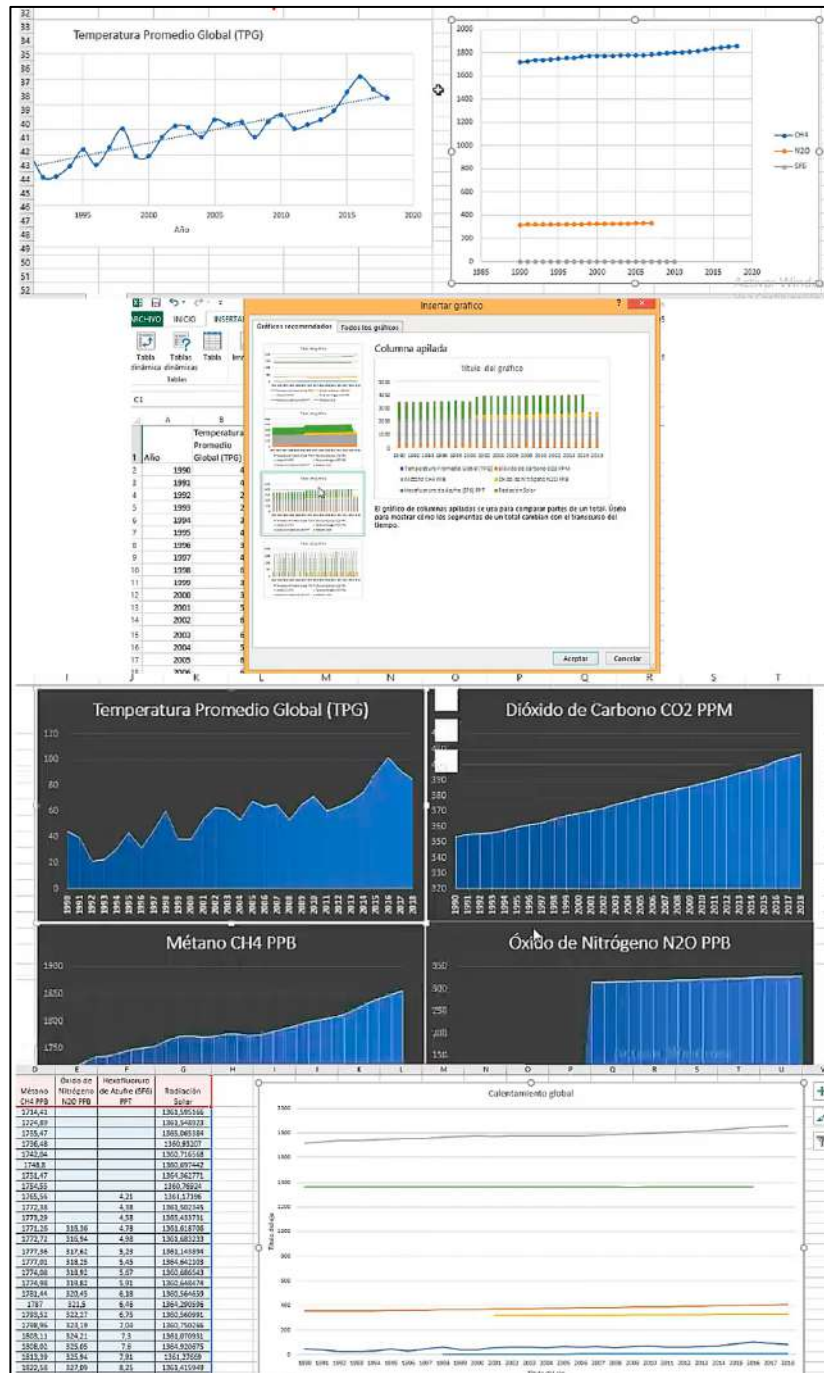


Figura 6. Primeros análisis de los estudiantes

Cambiaron entre una y otra representación gráfica, buscando la que les pareciera mejor. Algunos tomaron todo el conjunto de los datos y buscaron una de las

gráficas recomendadas por Excel para representarlos. Por supuesto, las representaciones que el software ofrecía combinaban los datos en una misma escala, dificultando su comparación (Figura 7).



Figura 7. Gráficos generados por Carolina seleccionando todos los datos

Cuando se les cuestionó sobre el uso de ciertos tipos de gráficos y la opción de usar otros, sus respuestas demostraban que no tenían un plan definido y no poseían suficientes conocimientos sobre el software con el que trabajarían; simplemente se aventuraron a tratar los datos con Excel. Carolina, por ejemplo, describe: “pues si... el problema es que acá... no sé si sea por tipo de Word (sic) que yo tengo, digamos si yo lo quiero cambiar por este de dispersión...entonces no me sale”.

Con base en estas primeras aproximaciones los estudiantes analizaron el fenómeno y trataron de describirlo. Las primeras conclusiones se basaron en la descripción del comportamiento de cada uno de los datos en función de los años. Prácticamente todos los estudiantes consideraron como variable en el eje horizontal de sus gráficos los años y en el eje vertical cada uno de los gases (o

todos en algunos casos). Esto hizo que sus conclusiones cayeran en el hecho de que todos aumentaban con el paso de los años, a excepción de los valores de la radiación solar. En el producto presentado por Lina y Maira, cuando trabajaron en pequeños grupos, describen estas relaciones de la siguiente manera:

“Se observa que desde el año 1990 hasta el año 2018 el gas de emisión de efecto invernadero (CO_2) aumentó cada año. Desde el año 1990 hasta el 2000 el gas aumentó. Desde el año 2000 hasta el 2005 el gas de efecto invernadero (CH_4) tuvo un disminución-aumento y aumento-disminución, pero, este no fue tan evidente. Desde 2005 hasta el 2018 aumentó. Este gas se empezó a evidenciar a partir del año 2001, lo cual, en tan poco tiempo con relación a los otros gases ha aumentado impresionantemente. Este gas se empezó a evidenciar a partir del año 1998, lo cual, aumenta cada vez más”.

El discurso de los estudiantes sugiere que no lograron distinguir entre el comportamiento de uno u otro gas comparándolos entre sí, ni tampoco identificaron la relación que tienen estos con la temperatura. En este primer ciclo de modelación los estudiantes se quedan en lo que denominan Lesh y Doerr (2003) matematización de los elementos involucrados, que corresponde a encontrar valores en la AGM y utilizarlos sin establecer una relación de estos con el fenómeno y de estos entre sí. Particularmente, varios estudiantes dejan de lado la variable fundamental del fenómeno, la temperatura promedio global.

Otros estudiantes buscaron la forma de concluir al respecto del fenómeno sin utilizar los datos empleados. Por ejemplo, Camilo escribe “[...] de acuerdo al Dr. Policarp Garay Martín [...] en su artículo [...] el calentamiento global se está dando debido al aumento de los niveles de CO_2 ”. En este mismo sentido María y Catalina explicaron de donde viene cada gas y qué efectos nocivos producen: “La mayor fuente de las emisiones de dióxido de carbono procede de la combustión del carbón, petróleo y gas de las centrales eléctricas, los automóviles y las instalaciones industriales”; “El metano no es tóxico. Su principal peligro para la salud son las quemaduras que puede provocar si entra en ignición”; “El dióxido de nitrógeno forma parte de un grupo de contaminantes gaseosos que se producen como consecuencia del tráfico rodado y de otros procesos de quema de combustibles fósiles como plantas de energía.”.

Esto muestra que su proceso de modelación aún no está lo suficientemente desarrollado como para cumplir con el objetivo, es decir, no han planteado un modelo que logre describir, explicar y predecir el fenómeno con base en los datos ofrecidos. En esta etapa, siete de los nueve estudiantes no consideraron el uso de un modelo matemático para el proceso, es más, ni siquiera sabían que en Excel se podían agregar ecuaciones a los gráficos.

Esto es coherente con el diseño las AGM, pues se busca el tránsito a través de los ciclos de modelación en los cuales las primeras percepciones, concepciones y planteamientos del estudiante se refinan y mejoran hasta que se consolidan los

modelos (Lesh y Doerr, 2003; Lehrer y English, 2018). De igual manera, los ciclos investigativos del PPDAC no están hechos para ser recorridos una única vez, los estudiantes abordan el problema con una perspectiva inicial, pero luego de interpretar, concluir y comunicar sus resultados, la revisión de estos los lleva a volver a abordar el problema, dominando mejor la dinámica del sistema, determinando qué variables no son significativas o están ahí para desviar la atención, trazando un plan mejor, descartando hipótesis y explorando los datos con nuevos análisis (Pfannkuch et al., 2018).

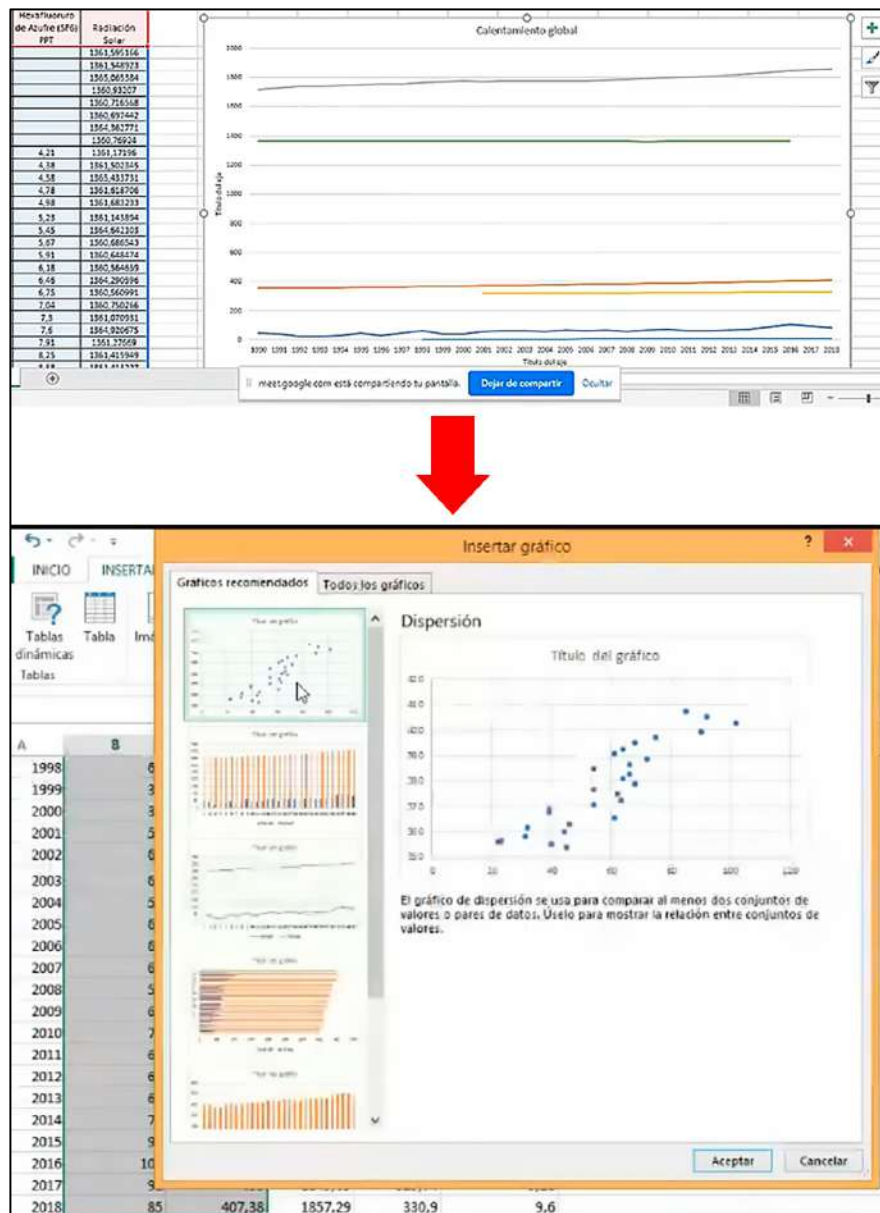


Figura 8. Refinamiento de la representación de datos

Refinamientos de los estudiantes

Un ejemplo de estos refinamientos que los estudiantes llevan a cabo se puede ver en la figura 8, donde la estudiante Catalina pasa de hacer una gráfica de líneas, que resulta confusa al seleccionar todo el conjunto de datos, hacia un gráfico de dispersión donde relaciona el comportamiento de la temperatura promedio global y los datos registrados de emisiones de metano.

Esto no solo es una mejora en cuanto a manejo y representación de datos, implica también una reflexión acerca de la utilidad del tipo de gráficas, ella escribe:

“Los datos bivariados se representan gráficamente por medio de diagramas de dispersión o de burbujas. Estos se utilizan para la interpretación de la relación entre las variables, pues se puede visualizar su intensidad, en función de la mayor o menor dispersión de la nube de puntos; su sentido, con respecto a si es directa o inversa; y su tipo, si es lineal o no, conforme a su tendencia.”

Varias muestras más de refinamiento se evidenciaron en la implementación de esta actividad. Carolina, por ejemplo, quien utilizó dos gráficos de barras (figura 7) para representar todos los datos (incluso temperatura versus temperatura estaba incluido), después de la revisión de sus conclusiones optó por usar un diagrama de dispersión para cada uno de los gases (figura 9) y mencionó que:

“Para que el análisis fuera más justo, lo ideal hubiera sido completar los datos faltantes de la tabla, sin embargo, en una exhaustiva búsqueda no encontré los datos globales, los puntuales, porque la IPCC arrojaba tablas a gran escala y, por otro lado, en la página de la Nasa tienen los ya determinados aquí”.

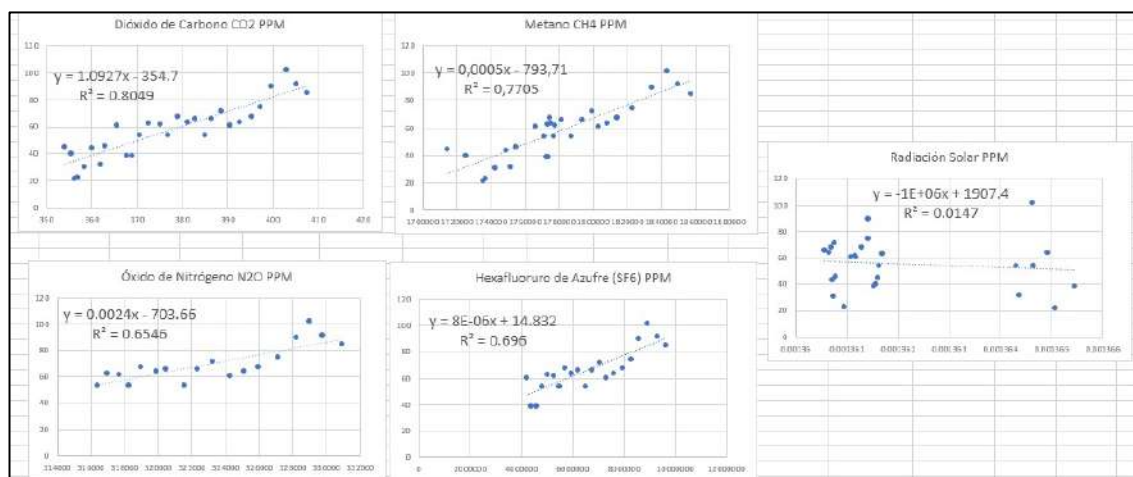


Figura 9. Gráficos de dispersión de Carolina

Además del uso correcto de las representaciones gráficas para describir el comportamiento de dos conjuntos de variables, los estudiantes obtuvieron los coeficientes de determinación, el modelo lineal de regresión y gráficas extras para

reforzar sus conclusiones. Estos elementos fueron agregados a los diagramas de dispersión, producto de las primeras muestras que hacían los estudiantes. Cuando uno de ellos presentó sus diagramas con modelos de regresión, los otros, atentos a esto, decidieron anexarlo, aun cuando no tenían idea del sentido que tenían estos en la descripción del fenómeno. Este proceder está documentado en Lesh y Doerr (2003), donde los estudiantes, apoyados entre ellos, ya sea en pequeños grupos o con el grupo en general, recogen ideas, elementos y métodos utilizados para mejorar sus trabajos.

Institucionalización del contenido

En la fase de diseño instruccional de institucionalización del contenido se trabajó con las rectas de regresión y se les dio la explicación de su uso e interpretación, la forma en la que se calcula y las herramientas alternativas de software que se pueden utilizar para otras situaciones. Los estudiantes, en sus trabajos finales, agregaron estos elementos a sus diagramas, los interpretaron y los usaron para describir el fenómeno. Sin embargo, no se quedaron solo en ello, nuevamente incluyeron gráficas de otros tipos para ahondar en el fenómeno.

Por ejemplo, en la figura 10 que fue incluida por Maira en su trabajo, se muestra el comportamiento de la temperatura en el siglo XX, resaltando los 10 años más cálidos y los 10 años más fríos, para reforzar el hecho de que el calentamiento global trae consigo temperaturas cada vez más altas. Si bien esto no hace parte de la generación de un modelo con base en el análisis estadístico de correlación y regresión lineal de los datos ofrecidos, permite complementar la descripción y predicción del fenómeno.

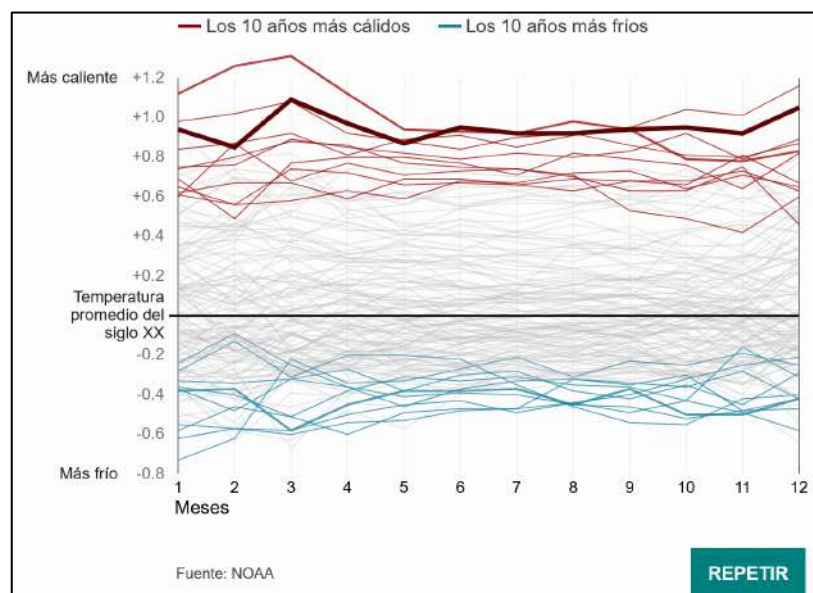


Figura 10. Gráfica anexada por Maira para reforzar sus conclusiones

También, los estudiantes usaron datos históricos y referenciaron análisis más profundos sobre las causas del fenómeno, escribieron sobre la revolución

industrial, por ejemplo, la quema de combustibles fósiles, la contaminación en general y la generación de emisiones producto de la ganadería extensiva. Todos estos elementos que relacionaron los participantes con el calentamiento global dan cuenta de la calidad de la actividad, pues, además de generar el modelo de regresión lineal buscado y calcular su coeficiente de determinación, se adentraron en el fenómeno; se comprometieron en el entendimiento de la situación descrita; documentaron sus procedimientos y centraron su labor en describir, explicar y predecir el comportamiento de dicho fenómeno. Teniendo en cuenta que es durante el tránsito a través de ciclos de modelación, en la refinación de sus modelos y en el mejoramiento de las conclusiones que realiza sobre la problemática, en donde el estudiante desarrolla sus conocimientos sobre el tema (Wild y Pfannkuch, 1999; Patel y Pfannkuch, 2018), es posible considerar a la AGM diseñada como una herramienta que, al ser implementada con el diseño instruccional descrito en este trabajo, puede garantizar el aprendizaje de la correlación y regresión lineal.

CONCLUSIONES

Del análisis del proceso de modelación estadística que realizan los estudiantes cuando desarrollan una AGM, fue posible concretar las características necesarias que deben poseer este tipo de herramientas para alcanzar el aprendizaje de la correlación y regresión lineal que se plantea como objetivo, a la vez que se identifican las limitaciones a superar, su pertinencia y la utilidad en entornos de aprendizaje virtuales, tanto a nivel pedagógico, como investigativo.

En específico, la AGM calentamiento global cumple a cabalidad los principios de diseño establecidos: realidad, construcción de modelos, documentación, autoevaluación, generalización y simplicidad. Promueve el tránsito de los estudiantes a través de los ciclos de modelación estadística y genera el espacio para la externalización de los procesos que realiza en su aprendizaje. Aún con las limitaciones metodológicas derivadas de la implementación virtual de las actividades, al desarrollar la AGM los participantes proporcionaron información suficiente para identificar la matematización de los elementos de la situación, las aproximaciones realizadas, sus refinamientos y la concreción de los modelos planteados. Esto la convierte en una herramienta que, además de garantizar el aprendizaje de la correlación y regresión lineal, resulta útil para la investigación en educación estadística, particularmente, dentro de la modelación estadística.

En este sentido, es posible esperar de los estudiantes, cuando desarrollan la AGM calentamiento global, que adquieran, amplíen y refinan sus conocimientos y habilidades estadísticas, mediante la organización, análisis y representación de datos, establecimiento de correlaciones, determinación de líneas de regresión y cálculo de coeficientes de ajuste. Además, con base en los modelos que ellos plantean, realizan conclusiones, explican, describen y predicen el comportamiento del fenómeno. Sin embargo, para que se dé este proceso de modelación, además

de la estructura y principios descritos, también se integra la presentación de la AGM como noticia en un portal simulado y las preguntas de ubicación, pues estos elementos cumplen el objetivo de involucrar al estudiante, lo cual, imbricado con la significancia de la situación trabajada (el calentamiento global), logra un amplio compromiso con el desarrollo completo de la actividad y el alcance de los objetivos.

Los modelos generados por los estudiantes fueron elaborados con base en el software Excel, mediante la herramienta de regresión lineal, en las nubes de puntos graficadas por ellos. Sin embargo, la concreción de la solución a la actividad no fue inmediata, como se esperaba, sus interpretaciones iniciales y primeras aproximaciones gráficas evidencian el refinamiento necesario y la importancia del uso de representaciones gráficas adecuadas para el logro del objetivo planteado. Además, la institucionalización del contenido y la enseñanza de la forma manual de cálculo de la recta de regresión, resaltan las bondades de la tecnología en el procesamiento de los datos estadísticos bivariados. Esto resulta relevante tanto para el estudiante como para el investigador, pues plantea un precedente en ambos sobre la necesidad de reflexión del problema antes de su abordaje, el tiempo necesario para el análisis inicial y los refinamientos, y la característica utilitaria de la tecnología en procedimientos estadísticos, siendo esta última relegada principalmente a la aritmética, el álgebra y el cálculo (López et al., 2019).

REFERENCIAS

- Abassian, A., Safi, F., Bush, S. y Bostic, J. (2020). Five different perspectives on mathematical modeling in mathematics education. *Investigations in Mathematics Learning*, 12(1), 53-65. <https://doi.org/10.1080/19477503.2019.1595360>
- Artigue, M., Douady, R., Moreno, L. y Gómez, P. (1995). *Ingeniería didáctica en educación matemática: un esquema para la investigación y la innovación en la enseñanza y el aprendizaje de las matemáticas*. Grupo Editorial Iberoamérica.
- Aridor, K. y Ben-Zvi, D. (2018). Statistical modeling to promote students' aggregate reasoning with sample and sampling. *ZDM-Mathematics Education*, 50(7), 1165-1181. <https://doi.org/10.1007/s11858-018-0994-5>
- Aridor, K. y Ben-Zvi, D. (2019). Students' aggregate reasoning with covariation. En G. Burril y D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research. ICME-13 monographs* (pp. 71-94). Springer. https://doi.org/10.1007/978-3-030-03472-6_4
- Blomhøj, M. (2009). Different perspectives in research on the teaching and learning of mathematical modelling – categorising the TSG21 papers. En M. Blomhøj y S. Carreira (Eds.), *Mathematical applications and modelling in the teaching and learning of mathematics*. pp. 1-17. Department of Science,

- Systems and Models, Roskilde University.
https://rucforsk.ruc.dk/ws/portalfiles/portal/3820977/IMFUFA_461.pdf#page=6
- Constantinou, C. P., Nicolaou, C. T. y Papaevripidou, M. (2019). A framework for modeling-based learning, teaching, and assessment. En A. Upmeyer zu Belzen, D. Krüger y J. van Driel (Eds.), *Towards a competence-based view on models and modeling in science education* (Vol. 12, pp. 39-58). Springer.
https://doi.org/10.1007/978-3-030-30255-9_3
- Doerr, H. y English, L. D. (2003). A modeling perspective on students' mathematical reasoning about data. *Journal for Research in Mathematics Education*, 34(2), 110-136. <https://doi.org/10.2307/30034902>
- Engel, J. y Sedlmeier, P. (2011). Correlation and regression in the training of teachers. En C. Batanero, G. Burril y C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (Vol. 14, pp. 247-258). Springer. https://doi.org/10.1007/978-94-007-1131-0_25
- Gea, M. (2014). *La correlación y regresión en bachillerato análisis de libros de texto y del conocimiento de los futuros profesores*. [Tesis de doctorado, Universidad de Granada]
<https://digibug.ugr.es/bitstream/handle/10481/34257/24031550.pdf?sequence=1>
- Gea, M., Batanero, C. y Roa, R. (2014). El sentido de la correlación y regresión. *Números*, 87, 25-35.
- Kaiser, G., Blomhøj, M. y Sriraman, B. (2006). Towards a didactical theory for mathematical modelling. *ZDM-Mathematics Education*, 38(2), 82-85.
<https://doi.org/10.1007/BF02655882>
- Lehrer, R. y English, E. (2018). Introducing children to modeling variability. En D. Ben-Zvi, K. Makar y J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 229-260). Springer. https://doi.org/10.1007/978-3-319-66195-7_7
- Lesh, R., Hoover, M. Hole, B., Kelly, A. y Post, T. (2000). Principles for developing thought-revealing activities for students and teachers. En A. Kelly y R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 591-646). Routledge.
- Lesh, R. y Doerr, H. (Eds.). (2003). *Beyond constructivism – Models and modeling perspectives on mathematics problem solving, learning and teaching*. Routledge. <https://doi.org/10.4324/9781410607713>
- Lesh, R. y Sriraman, B. (2010). Re-conceptualizing mathematics education as a design science. En B. Sriraman y L. English (Eds.), *Theories of mathematics education. Advances in Mathematics education* (pp. 123-146). Springer.
https://doi.org/10.1007/978-3-642-00742-2_14
- López, J. L., Pontes, A. y Varo, M. (2019). Las TIC en la enseñanza científico-técnica hispanoamericana: una revisión bibliográfica. *Digital Education Review*, 1(35), 229-243. <https://doi.org/10.1344/der.2019.35.229-243>

- Moliní, F. y Sánchez, D. (2019). Fomentar la participación en clase de los estudiantes universitarios y evaluarla. *REDU: revista de docencia universitaria*, 17(1), 211-227. <https://doi.org/10.4995/redu.2019.10702>
- Patel, A. y Pfannkuch, M. (2018). Developing a statistical modeling framework to characterize Year 7 students' reasoning. *ZDM-Mathematics Education*, 50(7), 1197-1212. <https://doi.org/10.1007/s11858-018-0960-2>
- Pfannkuch, M., Ben-Zvi, D. y Budgett, S. (2018). Innovations in statistical modeling to connect data, chance and context. *ZDM-Mathematics Education*, 50(7), 1113-1123. <https://doi.org/10.1007/s11858-018-0989-2>
- Seber, G. A. y Lee, A. J. (2003). *Linear regression análisis*. John Wiley y Sons. <https://doi.org/10.1002/9780471722199>
- Upmeier zu Belzen, A., van Driel, J. V. y Krüger, D. (2019). Introducing a framework for modeling competence. En A. Upmeier zu Belzen, D. Krüger y J. van Driel (Eds.) *Towards a competence-based view on models and modeling in science education* (pp. 3-19). Springer.
- Wild, C. J. y Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review*, 67(3), 223-248. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>
- Zieffler, A., Garfield, J. y Fry, E. (2018). What Is Statistics Education?. En D. Ben-Zvi, D. Makar y J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 37-70). Springer. https://doi.org/10.1007/978-3-319-66195-7_2

Elkin Alejandro Osorio Amaya
Fundación Universitaria Colombo
Germana, Colombia
docente.elkin.osorio@unigermana.edu.co

Santiago Inzunza Cázares
Universidad Autónoma de
Sinaloa, México
sinzunza@uas.edu.mx

Silvia Evelyn Ward Bringas
Universidad Autónoma de
Sinaloa, México
sevelynward@gmail.com

Recibido: febrero, 2022. Aceptado: diciembre, 2023
doi: 10.30827/pna.v17i3.23937



ISSN: 1887-3987

STATISTICAL MODELING FOR LEARNING CORRELATION AND LINEAR REGRESSION

Elkin Osorio Amaya, Santiago Inzunza Cáceres y Silvia Evelyn Ward

This paper shows the results of using a Model-Eliciting Activity (MEA) to learn correlation and linear regression. The MEA was designed based on real data on the phenomenon of global warming from official sources such as the National Aeronautics and Space Administration and the National Office of Oceanic and Atmospheric Administration. The activity was presented as a journalistic note on a news portal that provided students a set of data about different gas emissions over the years related to the phenomenon along with records of solar radiation levels. The gas data were given to the students with different concentration scales to promote in-depth comparative analyses; some data in different years were removed to motivate participants to complete them; and solar radiation was introduced to observe the existence of elements uncorrelated with the phenomenon but used to bias the analysis. With these elements, the proposed MEA meets the principles of design: reality, model building, documentation, self-assessment, generalization, and simplicity.

When solving the MEA, the participants acquired, expanded and refined their statistical knowledge and skills, organizing, analyzing and representing data, establishing correlations, determining regression lines and adjustment coefficients. Likewise, with the proposed models, they concluded, explained and described the phenomenon. It is possible to say that the MEA designed promotes the students' transit through the cycles of statistical modeling and generates the space for the externalization of the processes carried out in their learning.

Even with the methodological limitations derived from the virtual implementation of the activities, when developing the MEA the participants provided sufficient information to identify the mathematization of the elements of the situation, the approximations made, their refinements and the concretion of the proposed models. In addition to establishing the activity as a pedagogical tool and a useful element for research in virtual environments, the study allowed us to observe shortcomings in the students' modeling process, linked to planning prior to addressing the problem.