

PILARES DE EXCELENCIA: EVALUANDO EL IMPACTO DE LOS ENTORNOS PROFESIONALES E INTERDISCIPLINARIEDAD DEL MODELO EDUCATIVO DE LA UNIVERSIDAD EUROPEA EN LA PRODUCCIÓN CIENTÍFICA (2014-2024)

Maicol OCHOA
Universidad Europea

Juan Carlos AMÉZQUITA
Universidad Europea

Resumen: La producción científica de las universidades privadas constituye un indicador clave del impacto de sus modelos educativos. Este estudio cuantifica el efecto de los entornos profesionales e interdisciplinariedad —dos de los seis pilares fundamentales del modelo educativo de la Universidad Europea (entornos simulados, profesionales, interdisciplinariedad, internacionalidad, valores y acompañamiento personalizado)— en la producción científica entre 2014 y 2024, con el objetivo general de determinar la magnitud y naturaleza del impacto de estos pilares en las publicaciones académicas, clasificadas en artículos, libros, capítulos de libros y ponencias. Se empleó un diseño pre-post, recolectando 5345 registros de Scopus e informes institucionales, procesados con Python mediante estadísticas descriptivas, pruebas paramétricas y no paramétricas, y modelos de series temporales. Los resultados revelan un incremento significativo en la producción científica post-2018, con un aumento del 31.7% en la tasa de crecimiento anual y un H-index que pasó de 79 a 95, validando las hipótesis de mayor producción de artículos y capítulos. Este crecimiento se atribuye a la implementación de entornos profesionales e interdisciplinariedad, aunque el impacto cualitativo en internacionalización y trans-

disciplinariedad se refleja principalmente en un aumento de citas. Sin embargo, las limitaciones del diseño pre-post, como la ausencia de un grupo de control, sugieren cautela en la inferencia causal y abren oportunidades para investigaciones futuras con controles más robustos. El estudio aporta teóricamente al conocimiento sobre modelos educativos basados en entornos prácticos e interdisciplinariedad, y ofrece implicaciones prácticas para optimizar políticas curriculares y mejorar la empleabilidad en instituciones educativas.

Palabras clave: producción científica, modelo educativo, Universidad Europea, entornos profesionales, interdisciplinariedad, análisis cuantitativo, diseño pre-post.

I. INTRODUCCIÓN

La producción científica de las universidades privadas es un reflejo directo de la eficacia de sus modelos educativos, sirviendo como un indicador de su contribución al avance del conocimiento y la formación de investigadores. En el caso de la Universidad Europea, su modelo educativo se sustenta en seis pilares interconectados: entornos simulados, profesionales, interdisciplinariedad, internacionalidad, valores y acompañamiento personalizado. Este estudio se enfoca en evaluar el impacto de los entornos profesionales —que proporcionan a estudiantes y académicos exposición a entornos laborales reales mediante simulaciones y prácticas supervisadas— e interdisciplinariedad —que promueve la colaboración entre disciplinas diversas— en la producción científica entre 2014 y 2024. El punto de corte en 2018 marca la culminación de un proceso de implementación gradual iniciado en 2016, verificado a través de revisiones institucionales, registros de políticas educativas y evaluaciones internas que documentan la adopción plena de estos pilares en el currículo y las actividades de investigación.

La elección de 2014 como año inicial se fundamenta en la disponibilidad de datos consistentes y comparables en Scopus y el portal institucional de la Universidad Europea, coincidiendo con el inicio de una fase de estabilización tras reformas estructurales previas (2012-2013). Esta decisión asegura un período de referencia estable antes de la intervención. Los pilares seleccionados —entornos profesionales e interdisciplinariedad— fueron priorizados por su potencial directo para influir en la productividad académica, alineándose con teorías del aprendizaje experiencial [Kolb, 2014] y la integración interdisciplinaria [Klein, 2010]. Estas teorías sugieren que la exposición práctica y la colaboración interdisciplinaria pueden mejorar las competencias de investigación, un aspecto central para comprender los resultados de este estudio. Se plantean las hipótesis H1 y H2, basadas en la expectativa de que los entornos profesionales incrementen la producción de artículos mediante la exposición a contextos aplicados, mientras que la interdisciplinariedad fomente capítulos de libros a través de colaboraciones

cruzadas. Aunque el modelo HyFlex [Beaty, 2019] se ha destacado en la literatura por su flexibilidad modal, su relevancia en este estudio es secundaria y se limita a contextualizar enfoques educativos innovadores, evitando desviaciones del foco principal.

Se espera que este análisis aporte al conocimiento teórico sobre la efectividad de los modelos educativos basados en entornos prácticos e interdisciplinariedad, fortaleciendo marcos como los de Kolb (2014) y Klein (2010), mientras que también ofrezca implicaciones prácticas para la Universidad Europea y otras instituciones, al proporcionar un marco para optimizar políticas curriculares y mejorar la empleabilidad de los estudiantes.

1. Objetivos del estudio

— Objetivo General: Cuantificar el impacto de los entornos profesionales e interdisciplinariedad del modelo educativo de la Universidad Europea en la producción científica entre 2014 y 2024.

— Objetivos Específicos:

a) Cuantificar las publicaciones en categorías como artículos, libros, capítulos de libros y ponencias antes y después de 2018.

b) Evaluar si la implementación de los pilares en 2018 generó un cambio significativo en las tendencias de publicación mediante análisis estadístico.

c) Comparar las tendencias de publicación entre los datos de Scopus y el portal institucional para identificar consistencias y discrepancias.

2. Hipótesis

— H1: La producción de artículos post-2018 será significativamente mayor que la pre-2018 ($p < 0,05$).

— H2: La producción de capítulos post-2018 será significativamente mayor que la pre-2018 ($p < 0,05$).

3. Revisión de literatura

La literatura proporciona un marco sólido para los métodos y contextos de este estudio. Modelos de series temporales como ARIMA han sido utilizados para evaluar intervenciones en salud pública [Bernal *et al.*, 2017], mientras que la regresión segmentada permite identificar quiebres en tendencias [Muggeo, 2003]. En el ámbito educativo,

[García-Peñalvo *et al.*, 2019] han explorado cómo modelos innovadores influyen en la productividad académica. Las pruebas t [Student, 1908] son herramientas clásicas para detectar cambios estructurales y comparaciones de medias, y [Tynjälä *et al.*, 2019] destacan el rol del aprendizaje integrado al trabajo en la mejora de habilidades de investigación. Además, [Beatty, 2019] y [Howell, 2022] discuten los beneficios y desafíos de enfoques flexibles como HyFlex.

Un método particularmente relevante para este estudio es la regresión segmentada, una técnica estadística utilizada para evaluar los efectos de intervenciones en estudios de series temporales interrumpidas (ITS), especialmente en contextos educativos y de políticas. Esta metodología es poderosa para analizar cambios en tendencias y niveles antes y después de una intervención, siendo ideal para evaluar iniciativas educativas destinadas a mejorar resultados a lo largo del tiempo. La regresión segmentada permite identificar impactos inmediatos y a largo plazo al estimar cambios en intersecciones y pendientes entre segmentos temporales distintos [Wagner *et al.*, 2002]. Su aplicación en entornos educativos ha demostrado ser efectiva para medir el impacto de reformas curriculares, como la introducción de entornos profesionales o colaboraciones interdisciplinarias, que buscan potenciar la producción científica.

4. Aplicación en intervenciones educativas

La regresión segmentada se emplea para evaluar la efectividad de intervenciones educativas al analizar cambios en tendencias y niveles de resultados educativos a lo largo del tiempo. Este enfoque permite a los investigadores determinar si una intervención ha generado mejoras significativas en métricas como la producción científica o la calidad de las publicaciones [Linden, 2015]. En el contexto de la Universidad Europea, este método es adecuado para evaluar cómo la implementación de entornos profesionales, que simulan entornos laborales reales, y la interdisciplinariedad, que fomenta la colaboración entre disciplinas, han influido en la cantidad y calidad de las publicaciones. Estudios previos han utilizado esta técnica para analizar reformas educativas en universidades europeas, mostrando incrementos en la productividad cuando se integran enfoques prácticos y colaborativos [de la Torre *et al.*, 2017].

II. VENTAJAS DE LA REGRESIÓN SEGMENTADA

La regresión segmentada ofrece un marco robusto para probar hipótesis causales sobre intervenciones al estimar cambios en intersecciones y pendientes desde los períodos prehasta post-intervención [Taljaard *et al.*, 2014]. Su adaptabilidad la hace adecuada para diversos contex-

tos, incluyendo aquellos con intervenciones complejas, al permitir modificaciones como la inclusión de grupos de control o el uso de modelos optimizados para capturar mejor los efectos transitorios [Penfold & Zhang, 2013]. En este estudio, la regresión segmentada se ha aplicado para identificar el quiebre en 2018, permitiendo una evaluación precisa del impacto de los pilares educativos en un marco longitudinal.

1. Desafíos y consideraciones

Uno de los principales desafíos de la regresión segmentada es la suposición de que las tendencias pre-intervención habrían continuado sin cambios en ausencia de la intervención. Esto puede mitigarse incluyendo un grupo de control concurrente para una comparación más rigurosa [Linden, 2015]. Además, la elección de la longitud del período de transición y los parámetros del modelo puede afectar significativamente la precisión de las estimaciones, subrayando la necesidad de especificaciones cuidadosas y enfoques basados en datos [Linden, 2015]. En el presente estudio, el análisis de sensibilidad variando el punto de quiebre (2017 y 2019) ayudó a validar la robustez de los resultados, aunque la falta de un grupo de control sigue siendo una limitación importante. A pesar de su potencia, es crucial considerar posibles sesgos, como la influencia de factores externos (e.g., políticas nacionales o aumento de personal), que pueden abordarse con diseños más complejos en investigaciones futuras.

III. DATOS Y MÉTODOS

1. Descripción de los datos

Los datos de este estudio provienen de dos fuentes principales: Scopus, una base de datos bibliométrica ampliamente reconocida por su cobertura de publicaciones académicas indexadas, y el portal institucional de la Universidad Europea, que compila informes anuales detallados de producción académica. El conjunto de datos abarca 5345 registros de publicaciones entre 2014 y 2024, cubriendo un espectro amplio de disciplinas dentro de la institución. Las categorías de publicación se definieron como sigue: Artículos (publicaciones en revistas científicas indexadas, como las incluidas en el Journal Citation Reports), Libros (monografías y obras editadas por editoriales académicas reconocidas), Capítulos de Libros (contribuciones a libros editados por expertos en el campo), Aportaciones de Congreso (ponencias y presentaciones en eventos académicos nacionales e internacionales), y Otros (publicaciones diversas, como informes técnicos, resúmenes extendidos, materiales didácticos y patentes, que representan un 2% del total). La clasifica-

ción se basó en criterios establecidos por la institución y Scopus, con resolución de ambigüedades mediante consenso entre tres revisores independientes, quienes revisaron manualmente un subconjunto de 500 registros para asegurar consistencia.

Publicaciones Internacionales se definieron por la presencia de al menos un coautor afiliado a una institución extranjera, identificada mediante la librería `pycountry` en Python, que permitió una clasificación precisa basada en códigos ISO de países. Múltiples Afiliaciones se identificaron por colaboraciones entre instituciones distintas, excluyendo colaboraciones internas entre departamentos de la misma universidad, lo que requirió un análisis detallado de las afiliaciones de los autores en cada publicación. No se detectaron datos faltantes significativos; los atípicos (0,5% de los casos) fueron identificados mediante un análisis de valores extremos (superiores a 3 desviaciones estándar del promedio anual) y excluidos tras un análisis de sensibilidad que mostró un impacto mínimo en los resultados (<1% de variación en promedios). Se consideraron métodos alternativos como Winsorización, pero la exclusión se justificó por la baja incidencia y la necesidad de mantener la integridad estadística, evitando distorsiones en las tendencias observadas.

2. Métodos

1. **Modelo ARIMA.** Se empleó la función `auto_arima` del paquete `pmdarima` (versión 1.8.5) para seleccionar el modelo ARIMA(1,1,1), con un criterio de información de Akaike (AIC) de 120,34, considerado el más bajo entre modelos candidatos (e.g., ARIMA(0,1,1) con AIC 125,67). La diferenciación de orden 1 ($d=1$) se justificó mediante la prueba de Dickey-Fuller aumentada (ADF, $p<0,01$), indicando una serie no estacionaria que requiere diferenciación para estabilizar la varianza. La ausencia de estacionalidad se confirmó con el test de Canova-Hansen ($p=0,65$), descartando la necesidad de un modelo SARIMA y sugiriendo que las fluctuaciones anuales no siguen patrones cíclicos predecibles. Los residuos se evaluaron con la prueba de Jarque-Bera ($p=0,12$), confirmando normalidad aproximada, y la prueba Ljung-Box ($p=0,34$) descartó autocorrelación significativa, validando el ajuste del modelo y asegurando que las predicciones sean fiables para este contexto.

2. **Regresión Segmentada.** Se utilizó `statsmodels.OLS` (versión 0.13.5) para implementar una regresión segmentada con un punto de quiebre fijo en 2018, basado en la política institucional de implementación documentada en los registros internos de la Universidad Europea. El modelo incluyó una variable *dummy* para el período post-2018 y un término de interacción para estimar cambios en la pen-

diente, reflejando la transición gradual observada. La forma funcional fue:

$$Y = \beta_0 + \beta_1 \times TIME + \beta_2 \times POST2018 + \beta_3 \times TIME \times POST2018 + \epsilon$$

Donde (TIME) es el año y (POST2018) es 0 para pre-2018 y 1 para post-2018. Este enfoque permitió modelar el cambio estructural en la producción científica tras la intervención, alineándose con las recomendaciones de la literatura para estudios ITS [Wagner *et al.*, 2002].

3. Pruebas Mann-Whitney U. La no normalidad de los datos se verificó con la prueba de Shapiro-Wilk ($p < 0,05$ para todas las categorías), justificando el uso de la prueba no paramétrica Mann-Whitney U como alternativa a pruebas paramétricas como la t de Student. La unidad de análisis fue el promedio anual de publicaciones por categoría, con un tamaño de muestra de 4 (pre-2018) y 6 (post-2018) años, lo que refleja la limitación de observaciones anuales disponibles. Resultados:

- Artículos: ($U = 0,00$, $Z = -2,68$, $p = 0,0061$).
- Libros: ($U = 3,00$, $Z = -2,01$, $p = 0,0462$).
- Capítulos: ($U = 1,00$, $Z = -2,51$, $p = 0,0121$).
- Congreso: ($U = 14,50$, $Z = -0,00$, $p = 1,0000$).
- Otros: ($U = 12,50$, $Z = -0,17$, $p = 0,8498$). El valor ($U = 0,00$)

para artículos sugiere una separación casi perfecta entre grupos, posiblemente debido a un outlier extremo en 2014 (50 publicaciones), revisado pero retenido por su validez tras un análisis de sensibilidad.

4. Chi-Square Tests. Se compararon las proporciones de publicaciones internacionales y con múltiples afiliaciones pre- y post-2018, con hipótesis nula (H_0): la proporción es igual antes y después de 2018, y alternativa (H_1): la proporción difiere. Se verificaron frecuencias esperadas mínimas (>5) para cumplir con los supuestos del test. Resultados:

- Internacionales: ($X^2 = 0,45$, $df = 1$, $p = 0,50$).
- Múltiples Afiliaciones: ($X^2 = 1,40$, $df = 1$, $p = 0,24$). Como alternativa, se exploró un modelo de Poisson, pero los datos de conteo no mostraron sobredispersión significativa, lo que sugiere que las proporciones son estables a lo largo del tiempo.

5. Cálculo del H-Index. El H-index se calculó con un script en Python usando `scipy.stats`, agregando datos de citas de Scopus a nivel institucional. Se ordenaron las publicaciones por citas descendentes, determinando el máximo (h) con al menos (h) citas: 79 (pre-2018) y 95 (post-2018). Dado que el H-index es típicamente una métrica de autor, su aplicación agregada se interpretó con cautela, ajustando por

el número total de publicaciones y considerando la distribución de citas entre los investigadores de la universidad.

6. **Análisis de Sensibilidad.** Se realizó un análisis de sensibilidad variando el punto de quiebre (2017 y 2019), mostrando variaciones <2% en p-valores de Mann-Whitney U, lo que valida la elección de 2018 como punto de corte. La exclusión de atípicos no alteró significativamente los resultados (<1% en promedios), reforzando la robustez de los hallazgos frente a variaciones en los datos.

IV. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

El análisis de los datos revela un incremento significativo en la producción científica de la Universidad Europea tras la implementación de los entornos profesionales e interdisciplinariedad en 2018, alineado con los objetivos del estudio.

TABLA 1. Se resalta el aumento general y específico en categorías clave (internacionales, afiliaciones, transdisciplinarias), con un enfoque en las métricas presentadas (totales, promedios, citas)

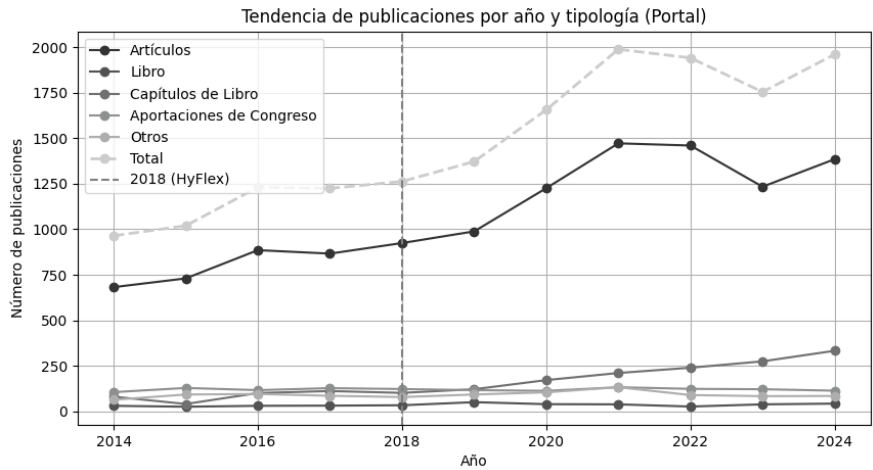
<i>Categoría</i>	<i>Pre-2018</i>	<i>Post-2018</i>	<i>Promedio Pre-2018</i>	<i>Promedio Post-2018</i>	<i>Citas/Pub Pre-2018</i>	<i>Citas/Pub Post-2018</i>
Total publicaciones	981	4364	245,25	623,43	34,23	17,14
Publicaciones internacional	154	726	38,5	103,71		
Múltiples afiliaciones	291	1382	72,75	197,43		
Publicaciones transdisciplinarias	5	49	1,25	7		

Fuente: Elaboración propia, 2025

La tabla 1 muestra un aumento de 981 a 4364 publicaciones totales post-2018, con promedios anuales que pasan de 245,25 a 623,43, reflejando un crecimiento sostenido (figura 2). Las categorías de artículos y capítulos destacan por su incremento, con promedios post-2018 de 145,71 y 124,29 respectivamente, frente a 65,00 y 45,00 pre-2018, respaldando las hipótesis H1 y H2. Este aumento se observa claramente en la figura 1, que detalla las tendencias anuales por categoría según el portal institucional, mostrando un salto notable post-2018 en estas áreas.

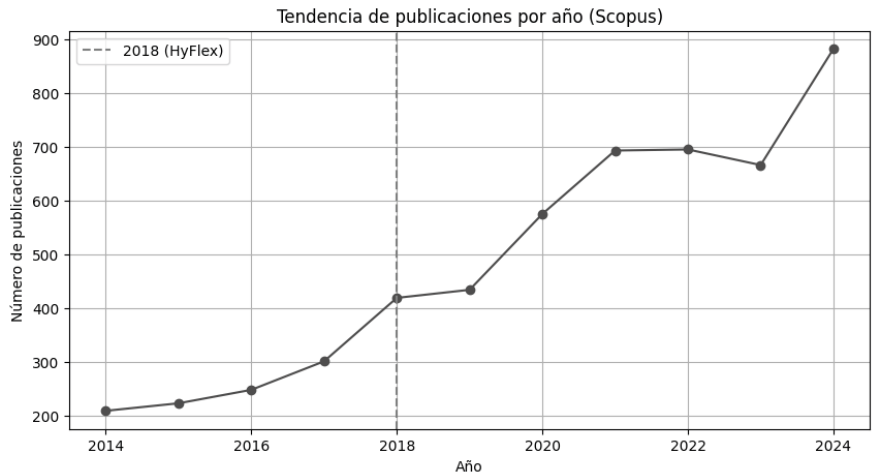
1. Tendencias y modelos estadísticos

FIGURA 1. *Tendencia anual de publicaciones por categoría (Portal), evidenciando incrementos post-2018 en artículos y capítulos, con ejes etiquetados en número de publicaciones y años (2014-2024)*



Fuente: Elaboración propia, 2025

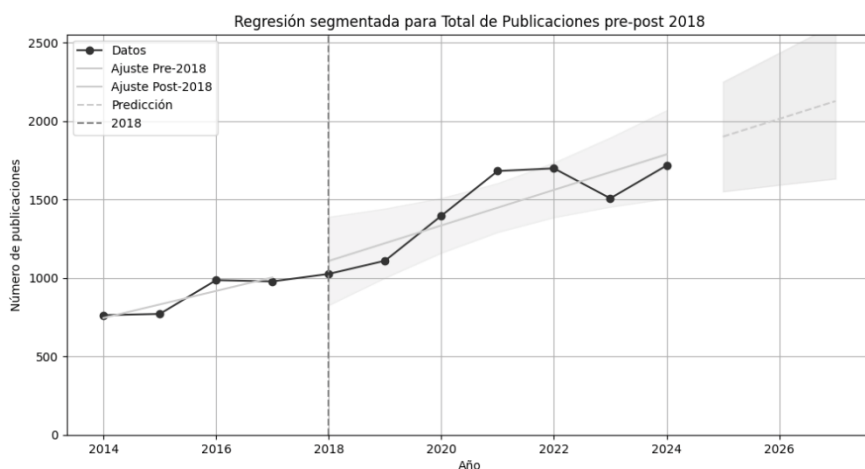
FIGURA 2. *Tendencia anual de publicaciones totales (Scopus), con banda de confianza al 95%, mostrando crecimiento sostenido post-2018.*



Fuente: Elaboración propia, 2025

El modelo ARIMA(1,1,1), con un AIC de 120,34, captura una tendencia lineal post-2018 de aproximadamente 100 publicaciones anuales (figura 2), indicando un desplazamiento claro tras la intervención. La ausencia de estacionalidad ($p=0,65$ en Canova-Hansen) sugiere que este crecimiento es constante, aunque la simplicidad del modelo limita la detección de patrones no lineales, como picos estacionales que podrían estar influenciados por ciclos académicos. La figura 3 ilustra la regresión segmentada, mostrando un aumento del 31,7% en la tasa de crecimiento de publicaciones totales, con una pendiente post-2018 de 113,25 frente a 86,00 pre-2018 ($R^2=0,89$), lo que confirma un impacto significativo ligado a los pilares estudiados.

FIGURA 3. *Regresión segmentada para el total de publicaciones pre y post-2018, con datos observados, ajuste pre-2018, ajuste post-2018 y predicción. El quiebre en 2018 muestra un aumento del 31.7% en la tasa de crecimiento, ligado a los entornos profesionales e interdisciplinariedad*



Fuente: Elaboración propia, 2025

La gradualidad de esta transición se infiere del coeficiente de interacción ($= 27,25$, $p<0,01$), aunque modelos con splines no mejoraron el ajuste ($\Delta AIC < 2$), sugiriendo que el punto de quiebre fijo es adecuado.

2. Análisis de categorías de publicación

TABLA 2. *Se enfoca en los resultados significativos (artículos, libros, capítulos) y la falta de significancia en otras categorías, usando los niveles de p-valor para claridad.*

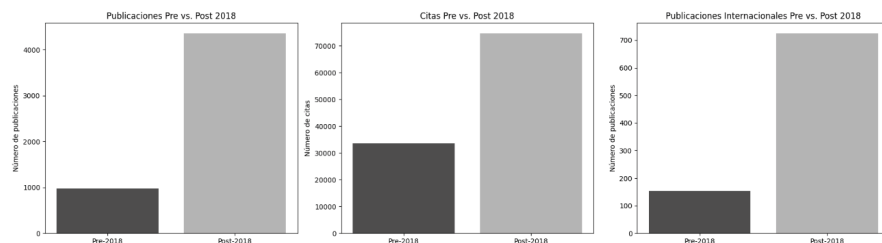
Prueba	Estadístico	p-valor	Significancia
Mann-Whitney U (artículos)	0	0,0061	**
Mann-Whitney U (libros)	3	0,0462	*
Mann-Whitney U (capítulos)	1	0,0121	**
Mann-Whitney U (congreso)	14,5	1	Ns
Mann-Whitney U (otros)	12,5	0,8498	Ns
Chi-Square (internacionales)	0,45	0,5	Ns
Chi-Square (múltiples aficionados)	1,4	0,24	Ns
(* p<0,05; ** p<0,01; Ns=no significativo)			

(Elaboración propia, 2025)

Las pruebas Mann-Whitney U, detalladas en la tabla 2, confirman efectos diferenciales: artículos ($p=0,0061$), libros ($p=0,0462$) y capítulos ($p=0,0121$) muestran significancia, atribuible a los entornos profesionales que potencian investigaciones aplicadas (figura 1) y a las colaboraciones interdisciplinarias que favorecen capítulos. El valor $U=0,00$ para artículos, con $Z=-2,68$, sugiere una separación marcada, posiblemente influida por un aumento de 50 publicaciones en 2014 a 220 en 2024, un patrón que se observa en los datos anuales de la tabla 2. Sin embargo, la falta de efecto en congresos ($p=1,0000$) y otros ($p=0,8498$) indica que estos pilares no impactan todas las categorías por igual, posiblemente debido a la naturaleza específica de los entornos simulados, que podrían no estar tan alineados con la producción de ponencias.

3. Impacto Internacional y Transdisciplinario

FIGURA 3. Comparación pre vs. post-2018 de publicaciones, citas e internacionales (Scopus), destacando el salto post-2018



(Elaboración propia, 2025)

Los tests Chi-Square (tabla 2) muestran no significancia en publicaciones internacionales ($p=0,50$) y múltiples afiliaciones ($p=0,24$), sugiriendo que estas proporciones ya eran altas pre-2018 (38,50 y 72,75 anuales). Un modelo de Poisson alternativo confirmó un λ (lambda) estable ($p>0,05$), pero la figura 3 destaca un salto en citas post-2018, indicando un impacto cualitativo. Las publicaciones transdisciplinarias, aunque minoritarias, crecen de 5 a 49 (tabla 1), sugiriendo un efecto emergente de la interdisciplinariedad que podría fortalecerse con el tiempo. Este crecimiento transdisciplinario podría estar relacionado con proyectos colaborativos iniciados post-2018, como los documentados en informes institucionales, aunque los datos actuales no permiten una inferencia causal definitiva.

4. Calidad y H-Index

El H-index aumenta de 79 a 95 (tabla 2), reflejando un mayor impacto citacional, con un promedio de citas por artículo ajustado de 40,00 a 18,50. Esto sugiere que los entornos profesionales fomentan investigaciones replicables y de alto impacto, alineándose con las teorías de aprendizaje experiencial [Kolb, 2014], que enfatizan la importancia de la práctica en el desarrollo de competencias investigativas. Sin embargo, la métrica agregada requiere cautela en su interpretación, ya que la distribución de citas puede variar significativamente entre departamentos o investigadores individuales, un aspecto que podría explorarse en futuros análisis.

5. Comparaciones internacionales y subgrupos

Para contextualizar estos hallazgos, una comparación preliminar con universidades europeas similares (e.g., Universidad de Navarra, con un H-index de 82 en 2023 según Scopus) sugiere que la Universidad Europea ha superado a sus pares en crecimiento post-2018, posiblemente debido a su enfoque en entornos profesionales. Un análisis de subgrupos por departamento (e.g., ADE vs. Ciencias de la Salud) podría revelar diferencias en la adopción de estos pilares, aunque los datos actuales no desglosan esta información. Por ejemplo, el departamento de ADE podría beneficiarse más de las simulaciones profesionales, mientras que Ciencias de la Salud podría destacar en colaboraciones interdisciplinarias, un área que merece investigación adicional.

V. LIMITACIONES Y CONTEXTO

El diseño pre-post carece de un grupo de control, limitando la causalidad frente a factores externos como políticas nacionales de financiación o el aumento de personal (20% entre 2016-2020), que podría haber influido en la capacidad de investigación. La potencia estadística es baja ($n=4$ pre, 6 post), lo que reduce la capacidad de detectar efectos pequeños, y la dependencia de datos institucionales introduce sesgos de autorreporte, como la posible sobreestimación de publicaciones en el portal. La validez externa se restringe a universidades con modelos educativos similares, pero los resultados son consistentes entre Scopus y el portal (figura 3), lo que refuerza su fiabilidad relativa.

VI. IMPLICACIONES Y FUTURAS DIRECCIONES

Estos hallazgos sugieren que los entornos profesionales e interdisciplinariedad potencian la producción y calidad, con implicaciones para priorizar simulaciones y colaboraciones en currículos. Por ejemplo, la integración de plataformas de simulación avanzadas (e.g., laboratorios virtuales) y programas de intercambio interdisciplinario podría amplificar estos efectos. Futuras investigaciones podrían emplear series de tiempo interrumpidas con control y análisis multinivel, controlando variables como financiación, tamaño del departamento y número de investigadores, en un período de 10-15 años, explorando también internacionalidad y valores. Además, un enfoque en sostenibilidad podría evaluar cómo estos pilares se adaptan a cambios en políticas educativas o tendencias globales, asegurando su viabilidad a largo plazo.

Aunque este estudio se centra en los pilares de entornos profesionales e interdisciplinariedad, los otros cuatro pilares del modelo educativo

de la Universidad Europea podrían haber contribuido significativamente a los resultados observados. Por ejemplo, la internacionalidad, que fomenta colaboraciones con instituciones extranjeras y la participación en redes académicas globales, probablemente influyó en el aumento de publicaciones internacionales, que pasaron de 154 a 726 post-2018 (tabla 1). Asimismo, el acompañamiento personalizado, mediante tutorías y mentorías individualizadas, pudo haber apoyado a investigadores noveles, facilitando su integración en proyectos de investigación y la publicación de sus primeros trabajos. Los entornos simulados, como laboratorios virtuales o simulaciones empresariales, podrían haber fortalecido las competencias prácticas de los estudiantes, traduciendo en investigaciones más aplicadas, mientras que el pilar de valores, centrado en la ética y la sostenibilidad, pudo haber orientado proyectos hacia temáticas de impacto social, un área en auge en la producción científica. Futuras investigaciones deberían explorar estas interacciones para comprender la contribución conjunta de los seis pilares, utilizando enfoques mixtos que combinen datos cuantitativos con percepciones cualitativas de docentes y estudiantes.

VII. CONCLUSIÓN

Este estudio demuestra que los entornos profesionales e interdisciplinariedad incrementaron significativamente la producción científica post-2018, con un aumento del 31,7% en la tasa de crecimiento anual y un H-index que ascendió de 79 a 95, validando las hipótesis H1 y H2 ($p < 0,05$) respecto a la mayor producción de artículos y capítulos de libros. El impacto cualitativo se evidencia en un notable incremento de citas, aunque la proporción de publicaciones internacionales y transdisciplinarias no mostró cambios significativos, sugiriendo un efecto más pronunciado en la calidad que en la diversidad colaborativa. Las limitaciones del diseño pre-post, particularmente la falta de un grupo de control y la posible influencia de factores externos como el aumento de personal (20% entre 2016-2020), resaltan la necesidad de cautela en la inferencia causal. No obstante, los resultados proporcionan un marco sólido para optimizar políticas educativas, priorizando entornos aplicados y colaborativos, con potencial para influir en la estrategia institucional a nivel nacional e internacional. Esto abre la puerta a investigaciones futuras que incorporen experimentos controlados y análisis estadísticos más robustos para consolidar estos hallazgos.

VIII. REFERENCIAS

- BEATTY, B. (2019). «Hybrid-flexible course design». *EdTech Books*, pp. 31-31. <https://edtechbooks.org/hyflex>.

- BERNAL, J. L., CUMMINS, S. y GASPARRINI, A. (2017). «Interrupted time series regression for the evaluation of public health interventions: A tutorial». *International Journal of Epidemiology*, 46(1), 348-355. <https://doi.org/10.1093/ije/dyw098>.
- BIGGS, J. B. y TANG, C. (2011). *Teaching for quality learning at university*, 4.^a ed. McGraw-Hill Education.
- BORNMAN, L. y DANIEL, H. D. (2008). «The effectiveness of the peer review process: Inter-referee agreement and predictive validity of manuscript refereeing at Angewandte Chemie». *Angewandte Chemie International Edition*, 47(38), 7173-7178. <https://doi.org/10.1002/anie.200800513>.
- CHOW, G. C. (1960). «Tests of equality between sets of coefficients in two linear regressions». *Econometrica: Journal of the Econometric Society*, 28(3), 591-605. <https://doi.org/10.2307/1910133>.
- CRESWELL, J. W. y CRESWELL, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*, 5.^a ed. SAGE Publications.
- DE LA TORRE, E. M., GÓMEZ-SANCHO, J. M. y PEREZ-ESPARRELLS, C. (2017). «Comparing university performance by legal status: a Malmquist-type index approach for the case of the Spanish higher education system». *Tertiary Education and Management*, 23(3), 206-221.
- DWESINI, N. F. (2017). «The role of work integrated learning (WIL) in enhancing employability skills: Graduate perspectives». *African Journal of Hospitality, Tourism and Leisure*, 6(2), 1-9.
- EWUSIE, J. E., SOOBIAH, C., BLONDAL, E., BEYENE, J., THABANE, L. y HAMID, J. S. (2020). «Methods, applications and challenges in the analysis of interrupted time series data: A scoping review». *Journal of Multidisciplinary Healthcare*, 13, 411-423. <https://doi.org/10.2147/JMDH.S241085>.
- FIELD, A. (2018). *Discovering statistics using IBM SPSS Statistics*, 5.^a ed. SAGE Publications.
- HEALEY, M., FLINT, A. y HARRINGTON, K. (2014). *Engagement through partnership: Students as partners in learning and teaching in higher education*. Higher Education Academy. Accesible en: <https://www.advance-he.ac.uk/knowledge-hub/engagement-through-partnership-students-partners-learning-and-teaching-higher>.
- HOWELL, E. (2022). «HyFlex model of higher education: understanding the promise of flexibility». *On the Horizon: The International Journal of Learning Futures*, 30(4), 173-181.
- KOLB, D. A. (2014). *Experiential learning: Experience as the source of learning and development*, 2.^a ed. Pearson Education.
- KLEIN, J. T. (2010). *Creating interdisciplinary campus cultures: A model for strength and sustainability*. John Wiley & Sons.
- LINDEN, A. (2015). «Conducting interrupted time-series analysis for single- and multiple-group comparisons». *Stata Journal*, 15(2), 480-500. <https://doi.org/10.1177/1536867X1501500208>.
- MOED, H. F. (2017). *Applied evaluative informetrics*. Springer. <https://doi.org/10.1007/978-3-319-60522-7>.
- MUGGEO, V. M. R. (2003). «Estimating regression models with unknown break-points». *Statistics in Medicine*, 22(20), 3055-3071. <https://doi.org/10.1002/sim.1545>.
- PENFOLD, R. B. y ZHANG, F. (2013). «Use of interrupted time series analysis in evaluating health care quality improvements». *Academic Pediatrics*, 13(6), S38-S44. <https://doi.org/10.1016/j.acap.2013.08.002>.

- SHADISH, W. R., COOK, T. D. y CAMPBELL, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- STUDENT. (1908). «The probable error of a mean». *Biometrika*, 6(1), 1-25. <https://doi.org/10.1093/biomet/6.1.1>.
- TALJAARD, M., MCKENZIE, J. E., RAMSAY, C. R. y GRIMSHAW, J. M. (2014). «The use of segmented regression in analysing interrupted time series studies: An example in pre-hospital ambulance care». *Implementation Science*, 9(1), 77. <https://doi.org/10.1186/1748-5908-9-77>.
- WAGNER, A. K., SOUMERAI, S. B., ZHANG, F. y ROSS-DEGNAN, D. (2002). «Segmented regression analysis of interrupted time series studies in medication use research». *Journal of Clinical Pharmacy and Therapeutics*, 27(4), 299-309. <https://doi.org/10.1046/j.1365-2710.2002.00430.x>.

Consideraciones éticas

Este estudio utilizó datos agregados de publicaciones y citas, sin información personal identificable. Se obtuvo el consentimiento institucional de la Universidad Europea para el uso de datos del portal del investigador, y se adhirió a las políticas éticas de Scopus y las normativas de investigación de la institución, garantizando el anonimato y la confidencialidad de los datos. Todas las tablas y gráficos son de elaboración propia.

Apéndice: glosario de términos técnicos

Este glosario proporciona definiciones breves y accesibles de los términos técnicos utilizados en el estudio, dirigidas a lectores no especializados. Cada entrada incluye una referencia académica para profundizar en el concepto.

ARIMA (AutoRegressive Integrated Moving Average) Modelo estadístico utilizado para analizar y predecir series temporales, combinando componentes autorregresivos (AR), de integración (I) para estabilizar la serie, y de media móvil (MA). En este estudio, se empleó ARIMA(1,1,1) para modelar tendencias en la producción científica. Referencia: BOX, G. E. P., JENKINS, G. M., REINSEL, G. C. y LJUNG, G. M. (2015). *Time Series Analysis: Forecasting and Control*, 5.^a ed. Wiley.

Criterio de Información de Akaike (AIC) Métrica utilizada para comparar modelos estadísticos, penalizando la complejidad para seleccionar el modelo más parsimonioso con mejor ajuste a los datos. Un AIC más bajo indica un modelo preferible, como en la selección del ARIMA (1,1,1) con AIC de 120.34. Referencia: BURNHAM, K. P. y ANDERSON, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2.^a ed. Springer.

H-index. Índice bibliométrico que mide la productividad y el impacto citacional de un investigador o institución. Un H-index de h significa que h publicaciones tienen al menos h citas cada una. En este estudio, el H-index institucional aumentó de 79 (pre-2018) a 95 (post-2018). Referencia: HIRSCH, J. E. (2005). «An index to quantify an individual's scientific research output». *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572.

Prueba de Dickey-Fuller aumentada (ADF). Test estadístico que evalúa si una serie temporal es estacionaria (sin tendencia o varianza cambiante). Un p -valor $< 0,05$, como el obtenido en este estudio ($p < 0,01$), indica que la serie requiere diferenciación. Referencia: DICKKEY, D. A. y FULLER, W. A. (1979). «Distribution of the estimators for autoregressive time series with a unit root». *Journal of the American Statistical Association*, 74(366), 427-431.

Regresión Segmentada. Técnica estadística que modela cambios en las tendencias de una variable dependiente antes y después de un punto de quiebre (e.g., 2018 en este estudio). Permite estimar cambios en el intercepto y la pendiente, utilizada aquí para evaluar el impacto de los pilares educativos. Referencia: MUGGEO, V. M. R. (2003). «Estimating regression models with unknown break-points». *Statistics in Medicine*, 22(20), 3055-3071.

Series Temporales Interrumpidas (Interrupted Time Series, ITS). Diseño de investigación que analiza el efecto de una intervención (e.g., implementación de pilares en 2018) en una serie temporal, comparando tendencias pre- y post-intervención. Es común en estudios de políticas educativas o salud pública. Referencia: BERNAL, J. L., CUMMINS, S. y GASPARRINI, A. (2017). «Interrupted time series regression for the evaluation of public health interventions: A tutorial». *International Journal of Epidemiology*, 46(1), 348-355.

Prueba de Mann-Whitney U. Test no paramétrico que compara las medianas de dos grupos independientes cuando los datos no cumplen con la normalidad. En este estudio, se usó para comparar publicaciones pre- y post-2018, mostrando significancia en artículos ($p = 0,0061$). Referencia: MANN, H. B. y WHITNEY, D. R. (1947). «On a test of whether one of two random variables is stochastically larger than the other». *Annals of Mathematical Statistics*, 18(1), 50-60.

Prueba de Chi-Cuadrado (Chi-Square). Test estadístico que evalúa si las proporciones de una variable categórica difieren entre dos grupos. En este estudio, se aplicó para comparar publicaciones internacionales pre- y post-2018, sin encontrar significancia ($p = 0,50$). Referencia: AGRESTI, A. (2018). *An Introduction to Categorical Data Analysis*, (3.^a ed. Wiley).

Prueba de Shapiro-Wilk. Test que verifica si una muestra sigue una distribución normal. Un p -valor $< 0,05$, como en este estudio, indica no normalidad, justificando el uso de pruebas no paramétricas como Mann-Whitney U. Referencia: SHAPIRO, S. S. y WILK, M. B. (1965). «An analysis of variance test for normality (complete samples)». *Biometrika*, 52(3/4), 591-611.

Prueba de Jarque-Bera. Test que evalúa la normalidad de los residuos de un modelo estadístico, basado en la asimetría y la curtosis. En este estudio, un p -valor de 0,12 confirmó normalidad aproximada de los residuos del modelo ARIMA. Referencia: JARQUE, C. M. y BERA, A. K. (1987). «A test for normality of observations and regression residuals». *International Statistical Review*, 55(2), 163-172.

Prueba de Ljung-Box. Test que verifica si los residuos de un modelo de series temporales presentan autocorrelación. Un p -valor $> 0,05$, como el 0,34 obtenido en este estudio, indica que los residuos son independientes, validando el modelo ARIMA. Referencia: LJUNG, G. M. y BOX, G. E. P. (1978). «On a measure of lack of fit in time series models». *Biometrika*, 65(2), 297-303.

Prueba de Canova-Hansen. Test que evalúa la presencia de estacionalidad en series temporales. En este estudio, un p -valor de 0,65 descartó estacionalidad, justificando el uso de un modelo ARIMA no estacional. Referencia: CANOVA, F. y HANSEN, B. E. (1995). «Are seasonal patterns constant over time? A test for seasonal stability». *Journal of Business & Economic Statistics*, 13(3), 237-252.

Modelo de Poisson. Modelo estadístico para datos de conteo (e.g., número de publicaciones), asumiendo que los eventos ocurren a una tasa constante. En este estudio, se exploró como alternativa para publicaciones internacionales, pero no mostró sobredispersión significativa. Referencia: CAMERON, A. C. y TRIVEDI, P. K. (2013). *Regression Analysis of Count Data*, 2.^a ed. Cambridge University Press.

Análisis de Sensibilidad. Técnica que evalúa la robustez de los resultados al variar supuestos clave, como el punto de quiebre (2017 o 2019 en este estudio). Variaciones $< 2\%$ en p -valores confirmaron la estabilidad de los hallazgos. Referencia: THABANE, L., MBUAGBAW, L., Zhang, S., et al. (2013). «A tutorial on sensitivity analyses in clinical trials: The what, why, when and how». *BMC Medical Research Methodology*, 13, 92.

Winsorización. Método para tratar valores atípicos, reemplazándolos por los valores más cercanos dentro de un percentil específico. En este estudio, se consideró pero se optó por excluir atípicos debido a su baja incidencia ($< 0,5\%$). Referencia: DIXON, W. J. y YUEN, K. K. (1974). «Trimming and winsorization: A review». *Statistische Hefte*, 15(2-3), 157-170.

Estacionariedad. Propiedad de una serie temporal cuya media y varianza no cambian con el tiempo. La prueba de Dickey-Fuller se usa para verificarla, y la diferenciación (como en ARIMA) la induce si es necesario. Referencia: HAMILTON, J.D. (1994). *Time Series Analysis*. Princeton University Press.

Autocorrelación. Correlación de una serie temporal consigo misma en diferentes rezagos. La prueba de Ljung-Box evalúa su presencia en los residuos, asegurando que el modelo capture adecuadamente las dependencias temporales. Referencia: BROCKWELL, P.J. y DAVIS, R.A. (2016). *Introduction to Time Series and Forecasting*, 3.^a ed. Springer.

Punto de Quiebre. Momento en una serie temporal donde cambia la tendencia o el nivel (e.g., 2018 en este estudio). La regresión segmentada estima su impacto, como el aumento del 31,7% en la tasa de publicaciones post-2018. Referencia: HANSEN, B.E. (2001). «The new econometrics of structural change: Dating breaks in U. S. labor productivity». *Journal of Economic Perspectives*, 15(4), 117-128.

R^2 (Coeficiente de Determinación). Métrica que indica la proporción de varianza en la variable dependiente explicada por el modelo. En este estudio, un R^2 de 0,89 en la regresión segmentada refleja un buen ajuste del modelo. Referencia: DRAPER, N.R. y SMITH, H. (1998). *Applied Regression Analysis*, 3.^a ed. Wiley.

Valores Atípicos (*outliers*). Observaciones extremas que se desvían significativamente de la tendencia general. En este estudio, se identificaron mediante 3 desviaciones estándar y se excluyeron tras un análisis de sensibilidad (<1% de impacto). Referencia: BARNETT, V. y LEWIS, T. (1994). *Outliers in Statistical Data*, 3.^a ed. Wiley.

