Se obtuvo una base de datos de Scopus con 2171 registros (de ahora en adelante *base cruda*), después de diferentes filtros y la construcción de un *theasure* obtuvimos una base de datos con 1069 registros. Primero, se eliminaron títulos duplicados con el paquete dplyr en R (Wickham *et al*., 2023). Para la elaboración del *theasure* y filtrar artículos poco relevantes, primero se obtuvo un set de palabras clave de los autores con una co-ocurrencia mayor a 10 en VOSview (van Eck & Waltman, 2010), se capturaron 109 palabras, las cuales se revisaron para decidir **qué** términos significaban los mismos y cuales no (ver tabla 1), así como cambiar el nombre común de las especies por los nombres científicos. La base anterior (*theasure*) y la *base cruda* se cargaron en R. A esta última se le aplico la función *tidyr:: separate\_rows* (Wickham *et al.,* 2024) que separó las palabras clave de los autores, subsecuentemente a las dos bases de datos se les elimino los espacios en blanco antes y después de cada palabra, caracteres alfanuméricos y se les convirtió en minúsculas. Este paso es necesario para homogeneizar las palabras y permitir que R pueda detectarlas correctamente.

Ahora bien, las palabas del *theasure* se usaron para modificar las palabras clave de los autores en la *base cruda*. Una vez se realizo el cambio, se filtraron todos aquellos registros que no tuvieran palabras clave de los autores y que no tuvieran al menos una palabra clave del *theasure*.

Adicionalmente, se realizó la depuración manual de artículos que sean revisiones bibliográficas, estén relacionados con la influencia de aves o mamíferos en la regeneración, economía sobre el manejo o rendimiento forestal, la influencia de la reproducción vegetativa o adición de hormonas en la regeneración, enfermedades (insectos, parásitos, hongos) que afecten la regeneración de las plántulas.

Con la base filtrada, se realizó un análisis de co-ocurrencia de las palabras clave de los autores en VOSview, el número de ocurrencias mínimas se especificó que fuera 10, dejándonos con 90 palabras clave de los autores, ademas, para la visualización se eliminaron las palabras usadas en la búsqueda bibliográfica. El análisis de co-ocurrencia se busca obtener una representación gráfica de la importancia entre los términos y como se agrupa la actividad de la investigación, así como identificar la tendencias en la investigación al gratificarnos el año promedio de cada palabra (van Eck & Waltman, 2010).

Con el paquete bibliometrix se obtuvo un listado de los 30 países con más artículos publicados y se graficaron los resultados con el paquete gglopt2 (Aria & Cuccurullo, 2017; Wickham, 2011). De igual forma, con bibliometrix, se realizó un análisis de estructura conceptual sobre los resumen, la función *conceptualStructure* primero realiza un analisis de correspondencia multiple para obtener la distancia Euclidania entre palabras, para después usar K-means y estimar lo grupos en los datos. Al realizar el análisis sobre el resumen y no las palabras clave podemos obtener un punto de vista diferente a la estructura cognitiva, ya que no solo se integran las palabras clave sino que también se integran palabras que faciliten la interpretación de los grupos formados. Se probaron diferentes parametros de la ocurrencia minima de las palabras y el numero de grupos, la mejor convinacion fue una co-ocurrencia minima de 85 y 10 grupos, esta convinacion nos permite capturar un numero alto de palabras, sin llegar a tener un exceso de las mismas (Aria & Cuccurullo, 2017).

Referencias

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of informetrics*, *11*(4), 959-975.

van Eck, N.J & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. **8**4, 523–538.

Wickham H, François R, Henry L, Müller K, Vaughan D (2023). dplyr: A Grammar of Data Manipulation. R package version 1.1.4, https://github.com/tidyverse/dplyr, https://dplyr.tidyverse.org.

Wickham H, Vaughan D, Girlich M (2024). tidyr: Tidy Messy Data. R package version 1.3.1, https://github.com/tidyverse/tidyr, [https://tidyr.tidyverse.org](https://tidyr.tidyverse.org/).

Wickham, H. (2011). ggplot2. *Wiley interdisciplinary reviews: computational statistics*, *3*(2), 180-185.