Title

Integration and synthesis of quantitative data: Alexander von Humboldt's renewed relevance in modern biogeography and ecology

Permalink

Journal

Authors

Keppel, Gunnar
Kreft, Holger

Publication Date

DOI

Copyright Information

OPINION

Integration and synthesis of quantitative data: Alexander von Humboldt's renewed relevance in modern biogeography and ecology

Gunnar Keppel[1,2,3,*] and Holger Kreft[1,4]

[1] Biodiversity, Macroecology & Biogeography, University of Goettingen, Büsgenweg 1, 37077 Göttingen, Germany.

[2] Natural and Built Environments Research Centre, School of Natural and Built Environments, University of South Australia, GPO Box 2471, SA 5001, Adelaide, Australia.

[3] Future Industries Institute, University of South Australia, GPO Box 2471, SA 5001, Adelaide, Australia.

[4] Centre of Biodiversity and Sustainable Land Use (CBL), University of Goettingen, Büsgenweg 1, 37077 Göttingen, Germany.

* corresponding author: gunnar.keppel@unisa.edu.au; http://orcid.org/0000-0001-7092-6149.

Abstract. The integration and synthesis of data from varied sources is becoming increasingly common in biogeography and ecology. Although analyses of large, global datasets from multiple, varied sources only became common recently, such studies trace their origin to the early 19th century. Alexander von Humboldt (1769-1859) was an eminent and much-celebrated scientist, who introduced the concept of collecting high-quality quantitative data across continents. More importantly, he championed inter-disciplinary research, data-driven studies, and global comparative syntheses. With these studies, he helped to lay the foundations for the fields of biogeography and ecology, and his ideas, concepts, and scientific approach remain at the forefront of exciting recent developments in both disciplines. Remembering Humboldt's 250th birthday, we show how his scientific philosophy and visionary approach to multi-disciplinary data integration and synthesis have recently been revitalized due to increasing computing power and the numerous complex environmental challenges facing our planet.

Keywords: Alexander von Humboldt, big data, data integration and synthesis, global change biology, macroecology, multi-disciplinarity

Introduction

Modern biogeography and ecology are increasingly reliant on integrating data from multiple, heterogeneous sources and across diverse disciplines (Hampton et al. 2013, Franklin et al. 2017, König et al. 2019). In the last decade, we have seen an explosive increase in studies focusing on large-scale data-driven syntheses. This increase in biogeographical and ecological studies using integrated global data has been facilitated by advances in computing, data storage, and analysis (Jetz et al. 2012, Kissling et al. 2018, König et al. 2019).

However, the origins of this approach date to Alexander von Humboldt (1769-1859) in the early 19th century. This remarkable scientist and science communicator was known for his paradigm-shifting intuitions, careful observations, and extensive and rigorous measurements and data recordings (Jackson 2009, Pausas and Bond 2019). As we are celebrating his 250th birthday this year, Humboldt's

way of conducting science remains pervasive and highly relevant for modern science (Kreft and Jetz 2007, Morueta-Holme and Svenning 2018, Pausas and Bond 2019). Here, we outline his legacy to the practice of modern biogeography and ecology, and we argue that he was the first to compile, integrate, and publish large quantitative datasets of diverse types.

Alexander von Humboldt

Alexander von Humboldt was born on 14 September 1769 in Berlin, then part of the Kingdom of Prussia. He displayed an early interest in nature and was nicknamed "the little pharmacist" at home due to his fondness for collecting and labelling insects, plants, and shells. His mother wanted him to become a civil servant, and he therefore undertook studies in commerce, foreign languages, geology, anatomy, and

1

astrology at various Prussian universities. In addition, he was able to develop his interest in nature through travels and meetings with naturalists Georg Forster and Sir Joseph Banks, who had both joined Captain James Cook on his expeditions to the Pacific (Botting 1974, Wulf 2015).

Humboldt is most famous for his expedition to the Americas. Equipped with the best and most precise available instruments, he travelled through Tenerife, Ecuador, Cuba, Venezuela, Mexico, and the United States of America for five years, collecting a broad spectrum of scientific data whenever and wherever he could (Dettelbach 1999, Wulf 2015). He aimed to "measure every possible quantity, 'armed' with every conceivable precise instrument from cyanometers to barometers" (Dettelbach 1999, pg. 480). His numerous collections, including some 6,000 plant specimens (Livi-Bacci 2015), and prolific writing of correspondence and subsequent publications (Schlegel 2006) gave this journey a long and far-reaching impact.

The influence of Alexander von Humboldt in science has been immense. He is perhaps best known for laying the foundations for the disciplines of biogeography and for providing the seeds for Charles Darwin's theory of evolution (Jackson 2009, Wulf 2015). However, Humboldt also made important contributions to other fields such as writing the first comprehensive socio-political history of Mexico (Livi-Bacci 2015), demonstrating that magnetic storms occur simultaneously in different parts of the planet (Malin and Barraclough 1991), and identifying the magnetic equator of Earth (Schlegel 2006).

One of his most enduring contributions was realizing and publicizing the intricate link between organisms and the environment (Morueta-Holme and Svenning 2018, Pausas and Bond 2019). This realization led him, among other things, to develop the concepts of tree lines, and altitudinal and latitudinal gradients of biodiversity (Nicolson 1987, Malanson et al. 2011, Moret et al. 2019), and to accurately predict global centers of plant diversity (Kreft and Jetz 2007). His detailed, data-rich records along altitudinal gradients were interdisciplinary and allowed him to document that species distributions corresponded to changes in physical parameters, such as temperature (Jackson 2009). The high precision of his records on the altitudinal distribution of vegetation on Mount Antisana allowed a recent study to document average upwards shifts of > 200 m in elevation for plant taxa (Moret et al. 2019).

However, it was the visionary way von Humboldt used his data and personal observations that has established him as one of the greatest naturalists and scientists. He was not content with simply presenting his vast, discipline-transcending data, but he integrated them into a more comprehensive picture of the functioning of our planet, bringing out the complex interactions between biosphere, atmosphere, and geosphere (Schlegel 2006, Moret et al. 2019). Humboldt also translated his findings into accessible and comprehensible publications and presentations (Dettelbach 1999, Jackson 2009). When Alexander von Humboldt passed away aged 89 in 1859, he was possibly the most famous scientist of his time, and his life was remembered and celebrated in Europe and the Americas for decades to come (Wulf 2015).

Humboldt and the integration and synthesis of data

Notwithstanding Humboldt's known legacy, here we would like to highlight an often overlooked contribution: his integrative and comprehensive approach to doing science. He was the first to practice a holistic, multi-disciplinary, and data-driven approach to science, which remains cutting edge to this day and has been referred to as "Humboldtian Science" (Cannon 1978). We believe that his approach involved three major paradigm shifts: (i) data collection strongly emphasized quantitative values, in addition to the then common practice of descriptive recordings, (ii) the scale of investigation was global, or as large as possible, instead of local, and (iii) analyses integrated as much data as possible instead of focusing on data collected in a single study.

As a result, Humboldt produced unprecedented numbers of records and measurements of climate, geography, geology, volcanism, and biodiversity (Botting 1974, Schlegel 2006, Livi-Bacci 2015). His integrated biotic and abiotic data along altitudinal gradients are the oldest available, and therefore important for documenting long-term altitudinal shifts under climate change (Morueta-Holme et al. 2015, Moret et al. 2019). He also covered an unprecedented geographic extent, collecting data from four continents to obtain holistic and broadly applicable insights (von Humboldt 1858, Morueta-Holme and Svenning 2018). Consequently, Humboldt's published works often transcended disciplines (von Humboldt and Bonplad 1805, 1858). Even today, collecting consistent detailed data on a (trans-) continental spatial scale remains a challenging and expensive, but worthwhile, exercise (Phillips and Miller 2002, Kao et al. 2012, Anderson-Teixeira et al. 2015). Humboldt achieved this before modern communication devices and air travel existed.

In addition, he actively sought out all available data by rigorously and persistently searching the literature and information from colleagues and repositories (Schlegel 2006). Humboldt therefore aspired to global networking of scientists and collaborations in an era when travel was expensive, slow and cumbersome, and communication took place via hand-written letters transported by horse and boat. Despite these obstacles, he established and maintained an extensive trans-disciplinary network of scientists, naturalists, politicians, artists, and intellectuals, creating huge volumes and a great variety of data at an unprecedented velocity. All data collected was carefully stored and filed and, where possible, confirmed with other experts. Andrea Wulf (2015) explains that "Humboldt collected his materials in boxes, which were divided by envelopes into different subjects" and that each envelope contained letters, "newspaper cuttings, pages from books, and pieces of paper on which he scribbled" (p. 236).

Not only did he collect, aggregate, store and manage these data, but he also developed new ways of analyzing and presenting them (Fig. 1). For example, he was the first to link plant distributions and vegetation features with physical data by supplementing an altitudinal profile of Chimborazo that displayed vegetation zones, with information on plant species, temperature, humidity, light refraction and intensity, and other features (Botting 1974, Jackson 2009). Humboldt also developed diagrams to compare his physical and biological data across mountains on different continents and invented isothermal lines to illustrate changes in temperature with altitude and latitude (von Humboldt and Bonplad 1805, Pausas and Bond 2019).

We therefore argue that Humboldt was the first person to face the challenges of big data. While the exact definitions vary among disciplines and authors, most agree that 'big data' involves unprecedented volumes and variety of data that accumulates at unprecedented rates (McAfee and Brynjolfsson
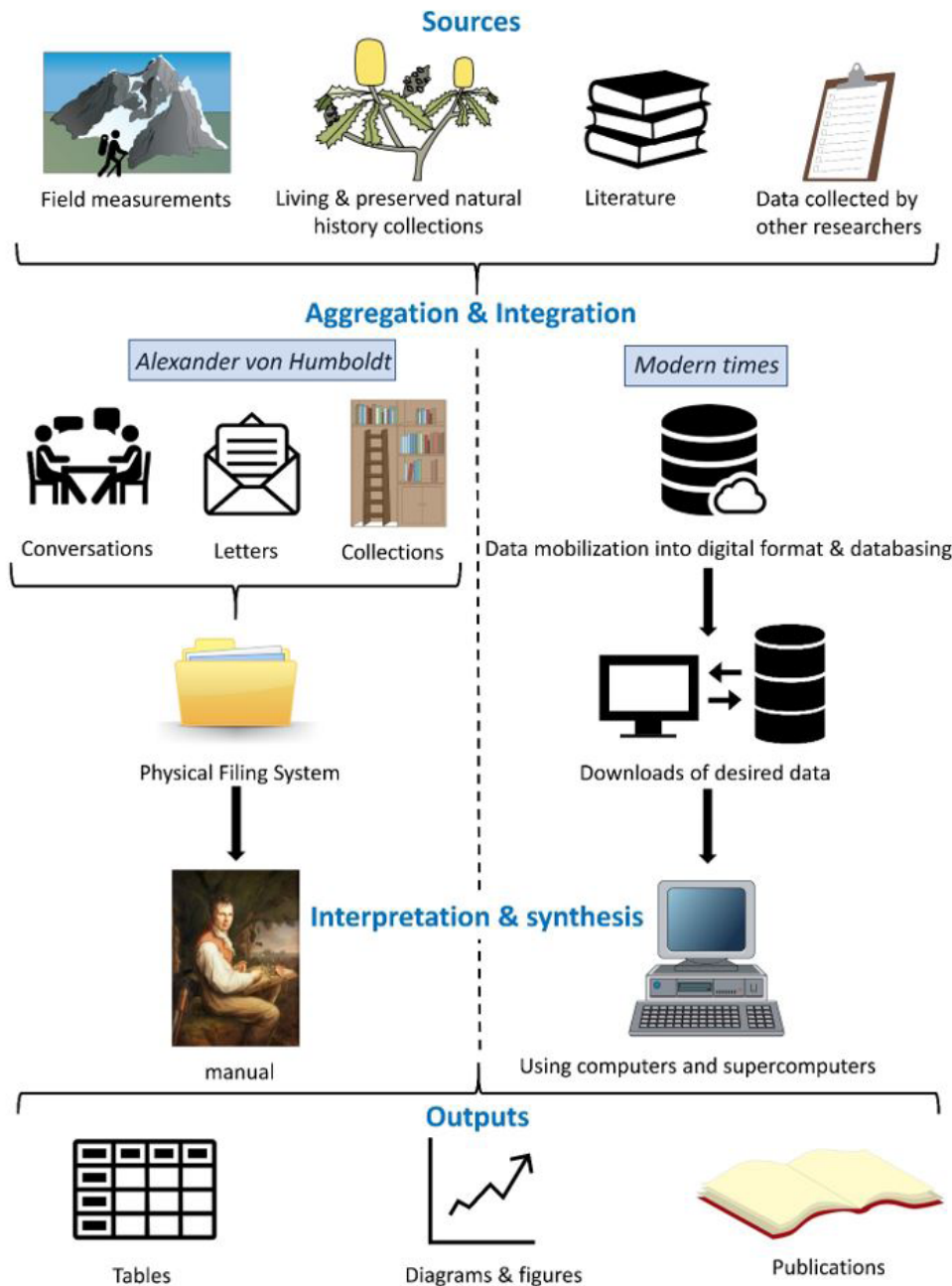
Figure 1. Data collection, integration, and synthesis as practiced by Alexander von Humboldt and in modern times. Note: pictures and diagrams taken from Ian Symbols (http://ian.umces.edu/symbols/) and Wikimedia Commons. Central photo an 1806 portrait of Alexander von Humboldt by Friedrich Georg Weitsch and the diagram of Mt. Chimborazo from *Essai sur la géographie des plantes* (von Humboldt and Bonplad 1805).

2012, Peters et al. 2014). Big data therefore also requires innovative forms of information processing and data presentation. As outlined above, Humboldt did face and overcome such challenges. However, current practices of working with large, global and interdisciplinary data are obviously different. Much more and a much wider range of data are available as a result of 200 years of scientific research, and an increasing amount of these data are digitally available (Kissling et al. 2018, König et al. 2019). Consequently, most recent studies with a global and interdisciplinary focus extract information from several databases to address their research questions (Fig. 1). The aggregation of increasingly large and heterogenous data, and subsequent analyses and presentation, are facilitated by rapidly advancing information technology (Jetz et al. 2012, König et al. 2019).

Recent renaissance of Humboldt's scientific approach

Following a trend of increased specialization of biological sciences into various disciplines through most of the 20th century, Alexander von Humboldt's vision for global, collaborative science based on comprehensive data has returned to the forefront of modern biology in recent decades (Kreft and Jetz 2007, Hampton et al. 2013, Balvanera et al. 2014, Kissling et al. 2018). Advances in computing, data storage and analysis, and communication have facilitated data integration at unprecedented scales (Jetz et al. 2012, Hallgren et al. 2016, Kissling et al. 2018). While standardized inter-continental data collection was a revolutionary idea at the time, it is now common in our efforts to gain a holistic understanding of our planet (Anderson-Teixeira et al. 2015) and to study anthropogenic climate change (IPCC 2014). As a result, analyses based on large, integrated datasets are becoming increasingly common in biogeography and ecology (Hampton et al. 2013, Franklin et al. 2017), using the same general approach first championed by Humboldt (Fig. 1).

Humboldt believed that the integration of knowledge from diverse disciplines is essential for understanding complex socio-ecological systems and included comparisons between natural and anthropogenic landscapes in his various discussions and diagrams (von Humboldt and Bonpland 1805). This insight remains highly relevant in modern conservation science because the consequences of human actions are having increasingly complex impacts on our planet that can only be explained and forecast using comprehensive data from multiple disciplines (IPCC 2014, Franklin et al. 2017, Ripple et al. 2017, IPBES 2019). Because global environmental change is increasingly posing existential threats to humanity (Ripple et al. 2017), international collaboration, data sharing, and interdisciplinary research are proliferating at an unprecedented scale in global change science (Hansen et al. 2013, IPCC 2014, IPBES 2019). The increased collaboration and rapidly evolving methods for data integration and analysis are producing important results, such as

facilitating increasingly accurate predictions of future environmental conditions and the associated impacts on biodiversity (Schuur et al. 2015, Urban et al. 2016).

In addition, remote sensing and online data repositories have prompted many global-scale studies incorporating biological, geographical, and climate data (e.g., Hansen et al. 2013, Morueta-Holme and Svenning 2018, Weigelt et al. 2019). Therefore, Humboldt's vision of global, multi-disciplinary data integration and synthesis is being increasingly realized. However, while modern information technology allows accessing and processing much more data than during the time of von Humboldt, the aggregation and analysis of large datasets can also result in a reduction of data quality and in flawed analyses (Meyer et al. 2016, Franklin et al. 2017), as data sources may be biased (Meyer et al. 2016) and can contain numerous errors, if not carefully curated (Mesibov 2013).

Many of the ideas and concepts developed by Alexander von Humboldt therefore continue to be at the forefront of some of the most exciting recent developments in ecology and biogeography. Humboldt's approach to science, especially with respect to using collaborative, inter-disciplinary, big data approaches to investigate phenomena, has become the cornerstone of modern science. This is exemplified in global change science, where an inter-disciplinary, collaborative approach is helping us better understand our planet and the effects humans are having (IPCC 2014). As the impacts of climate change will become more acute, and the availability of and processing power for big data will increase, the impact of the work and scientific philosophy of Alexander von Humboldt will remain relevant over decades to come.

Acknowledgments

References

Anderson-Teixeira, K.J., Davies, S.J., Bennett, A.C., et al. (2015) CTFS-ForestGEO: a worldwide network monitoring forests in an era of global change. Global Change Biology, 21, 528–549.

Balvanera, P., Siddique. I., Dee, L., Paquette, A., Isbell, F., Gonzalez, A., Byrnes, J., O'Connor, M.I., Hungate, B.A. & Griffin, J.N. (2014) Linking biodiversity and ecosystem services: current uncertainties and the necessary next steps. Bioscience, 64, 49–57.

Botting, D. (1974) Alexander von Humboldt. Prestel Verlag, Munich.

Cannon, S.F. (1978) Science in culture: the early Victorian period. Science History Publications, New York.

Dettelbach, M. (1999) The face of nature: precise measurement, mapping, and sensibility in the work of Alexander von Humboldt. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 30, 473–504.

Franklin, J., Serra-Diaz, J.M., Syphard, A.D. & Regan, H.M. (2017) Big data for forecasting the impacts of global change on plant communities. Global Ecology and Biogeography, 26, 6–17.

Hallgren, W., Beaumont, L., Bowness, A., et al. (2016) The biodiversity and climate change virtual laboratory: where ecology meets big data. Environmental Modelling & Software, 76, 182–186.

Hampton, S.E., Strasser, C.A., Tewksbury, J.J., Gram, W.K., Budden, A.E., Batcheller, A.L., Duke, C.S. & Porter, J.H. (2013) Big data and the future of ecology. Frontiers in Ecology and the Environment, 11, 156–162.

Hansen, M.C., Potapov, P.V., Moore, R., et al. (2013) High-resolution global maps of 21st-century forest cover change. Science, 342, 850–853.

IPBES (2019) Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Advanced unedited version, 6 May 2019. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Paris.

IPCC (2014) Synthesis report: climate change 2014. Intergovernmental Panel on Climate Change, Geneva.

Jackson, S.T. (2009) Alexander von Humboldt and the general physics of the Earth. Science, 324, 596–597.

Jetz, W., McPherson, J.M. & Guralnick, R.P. (2012) Integrating biodiversity distribution knowledge: toward a global map of life. Trends in Ecology & Evolution, 27, 151–159.

Kao, R.H., Gibson, C.M., Gallery, R.E., et al. (2012) NEON terrestrial field observations: designing continental-scale, standardized sampling. Ecosphere 3, art115.

Kissling, W.D., Walls, R., Bowser, A., et al. (2018) Towards global data products of Essential Biodiversity Variables on species traits. Nature Ecology & Evolution, 2, 1531–1540.

König, C., Weigelt, P., Schrader, J., Taylor, A., Kattge, J. & Kreft, H. (2019) Biodiversity data integration -

the significance of data resolution and domain. PLoS Biology, 17, e3000183.

Kreft, H. & Jetz, W. (2007) Global patterns and determinants of vascular plant diversity. Proceedings of the National Academy of Sciences of the USA, 104, 5925–5930.

Livi-Bacci, M. (2015) A note on Alexander von Humboldt. Population and Development Review, 41, 537–540.

Malanson, G.P., Resler, L.M., Bader, M.Y., Holtmeier, F.-K., Butler, D.R., Weiss, D.J., Daniels, L.D. & Fagre, D.B. (2011). Mountain treelines: a roadmap for research orientation. Arctic, Antarctic, and Alpine Research, 43, 167–177.

Malin, S.R.C. & Barraclough, D.R. (1991) Humboldt and the Earth's magnetic field. Quarterly Journal of the Royal Astronomical Society, 32, 279–293.

McAfee, A. & Brynjolfsson, E. (2012) Big data: the management revolution. Harvard Business Review, 90, 60–68.

Mesibov, R. (2013) A specialist's audit of aggregated occurrence records. ZooKeys, 293, 1–18.

Meyer, C., Weigelt, P. & Kreft, H. (2016) Multidimensional biases, gaps and uncertainties in global plant occurrence information. Ecology Letters, 19, 992–1006.

Moret, P., Muriel, P., Jaramillo, R. & Dangles, O. (2019) Humboldt's *Tableau Physique* revisited. Proceedings of the National Academy of Sciences of the USA, doi:10.1073/pnas.1904585116.

Morueta-Holme, N., Engemann, K., Sandoval-Acuña, P., Jonas, J.D., Segnitz, R.M. & Svenning, J.-C. (2015) Strong upslope shifts in Chimborazo's vegetation over two centuries since Humboldt. Proceedings of the National Academy of Sciences of the USA, 112, 12741–12745.

Morueta-Holme, N. & Svenning, J.-C. (2018) Geography of plants in the New World: Humboldt's relevance in the age of big data. Annals of the Missouri Botanical Garden, 103, 315–329.

Nicolson, M. (1987) Alexander von Humboldt, Humboldtian science and the origins of the study of vegetation. History of Science, 25, 167–194.

Pausas, J.G. & Bond, W.J. (2019) Humboldt and the reinvention of nature. Journal of Ecology, 107, 1031-1037.

Peters, D.P.C., Havstad, K.M., Cushing, J., Tweedie, C., Fuentes, O. & Villanueva-Rosales, N. (2014) Harnessing the power of big data: infusing

the scientific method with machine learning to transform ecology. Ecosphere, 5, art67.

Phillips, O. & Miller, J.S. (2002) Global patterns of plant diversity: Alwyn H. Gentry's forest transect data set. Missouri Botanical Press, St. Louis, MO.

Ripple, W.J., Wolf, C., Newsome, T.M., Galetti, M., Alamgir, M., Crist, E., Mahmoud, M.I. & Laurance, W.F. (2017) World scientists' warning to humanity: a second notice. Bioscience, 67, 1026–1028.

Schlegel, K. (2006) Space weather and Alexander von Humboldt's Kosmos. Space Weather, 4, doi:10.1029/2005SW000166.

Schuur, E.A.G., McGuire, A.D., Schädel, C., et al. (2015) Climate change and the permafrost carbon feedback. Nature, 520, 171–179.

Urban, M.C., Bocedi, G., Hendry, A.P., et al. (2016) Improving the forecast for biodiversity under climate change. Science, 353, aad8466.

von Humboldt, A. (1858) Cosmos: a sketch of a physical description of the universe. Harper & Brothers Publishers, New York.

von Humboldt, A. & Bonplad, A. (1805) Essai sur la géographie des plantes. Chez Levrault, Schoell et Compagnie, Libraires, Paris.

Weigelt, P., C. König & Kreft, H. (2019) GIFT – A Global Inventory of Floras and Traits for macroecology and biogeography. Journal of Biogeography doi:10.1101/535005.

Wulf, A. (2015) The invention of nature. The adventures of Alexander von Humboldt, the lost hero of science. John Murray, London.