Journal of Vegetation Science **22** (2011) 577–581

SPECIAL FEATURE: ECOINFORMATICS

Ecoinformatics and global change – an overdue liaison

Jürgen Dengler, Jörg Ewald, Ingolf Kühn & Robert K. Peet
SPECIAL FEATURE EDITORS

Keywords
Biodiversity; Bioinformatics; Data sharing; Database; Macroecology; Modelling; Phytosociology; Species distribution model; Vegetation plot.

Abbreviations
GBIF = Global Biodiversity Information Facility

Dengler, J. (corresponding author, dengler@botanik.uni-hamburg.de): Biodiversity, Evolution and Ecology of Plants, Biocentre Klein Flottbek and Botanical Garden, University of Hamburg, Ohnhorststr. 18, 22609 Hamburg, Germany
Ewald, J. (joerg.ewald@hswt.de): Botany & Vegetation Science, Faculty of Forestry, University of Applied Sciences Weihenstephan-Triesdorf, Hans-Carl-von-Carlowitz-Platz 3, 85354 Freising, Germany
Kühn, I. (ingolf.kuehn@ufz.de): Helmholtz Centre for Environmental Research GmbH – UFZ, Dept. Community Ecology, Theodor-Lieser-Str. 4, 06120 Halle, Germany
Peet, R.K. (peet@unc.edu): Department of Biology CB#3280, University of North Carolina, Chapel Hill, NC 27599-3280, United States.

Abstract

The field of ecoinformatics provides concepts, methods and standards to guide management and analysis of ecological data with particular emphasis on exploration of co-occurrences of organisms and their linkage to environmental conditions and taxon attributes. In this editorial, introducing the Special Feature 'Ecoinformatics and global change', we reflect on the development of ecoinformatics and explore its importance for future global change research with special focus on vegetation-plot data. We show how papers in this Special Feature illustrate important directions and approaches in this emerging field. We suggest that ecoinformatics has the potential to make profound contributions to pure and applied sciences, and that the analyses, databases, meta-databases, data exchange formats and analytical tools presented in this Special Feature advance this approach to vegetation science and illustrate and address important open questions. We conclude by describing important future directions for the development of the field including incentives for data sharing, creation of tools for more robust statistical analysis, utilities for integration of data that conform to divergent taxonomic standards, and databases that provide detailed plot-specific data so as to allow users to find and access data appropriate to their research needs.

Introduction

Human-induced global change might have an unprecedentedly strong negative impact on the biodiversity of the earth. Vitousek (1994) identified increased atmospheric CO_2, alteration of the biogeochemistry of the global nitrogen cycle and ongoing land-use and land-cover changes as the most important causes. Sala et al. (2000) agreed, adding climate change and biotic exchange to the list. While we have good data for the physical-chemical drivers of global change (i.e., climate change, nitrogen deposition, acid rain, increased CO_2), and the knowledge of current land use and its changes is reasonable at the global scale, quantitative data on the current state of biodiversity, its trends, and the way it is influenced by the different components of global change and their interactions (including biotic exchanges) are difficult to obtain. Despite the fact that some industrialised countries have set up elaborate biodiversity monitoring schemes during recent years (e.g. BDM in Switzerland, Hintermann et al. 2000), comparable data from the past are missing. At the global scale, data coverage is even more sparse, and analysing available data is hampered by a multitude of underlying sampling approaches and data-quality issues. Here ecoinformatics (i.e., the science of storing, retrieving and integrating ecological datasets; including biodiversity informatics) could provide important tools for collecting biodiversity and ecological data from many different sources and combining and analysing them in order to efficiently detect, understand, forecast and ultimately counteract changes of biodiversity (e.g. Bisby 2000; Canhos et al. 2004; Soberón & Peterson

Doi: 10.1111/j.1654-1103.2011.01313.x

2004; Jones et al. 2006; Bekker et al. 2007; Guralnick & Hill 2009; Recknagel 2011).

This Special Feature of the *Journal of Vegetation Science* contains contributions from the 9th International Meeting on Vegetation Databases held 24–26 February 2010 in Hamburg, Germany. Meetings on Vegetation Databases have been organised by the German Working Group on Vegetation Databases (now a Section within the German Network for Phytodiversity – NetPhyD) since 2002. With a focus on *Vegetation databases and climate change*, the conference in Hamburg, which was co-organised by the IAVS Working Group for Ecoinformatics, attracted 142 participants from 28 countries. It was the largest and most international within this series, and one of the largest ecoinformatics conferences ever held (for a detailed report, see Ewald et al. 2010).

In this Special Feature we present 14 selected papers from the Hamburg meeting, which cover a wide range of topics relevant to ecoinformatics, from databases and software tools through analytical and modelling approaches to applications, all with specific focus on vegetation data. A companion Special Volume from the conference under the title *Vegetation databases for the 21st century* will appear in *Biodiversity & Ecology*, comprising mainly *Database Reports* that describe the wealth of existing vegetation databases worldwide and their applications (Dengler et al. in press).

History of Ecoinformatics

'Ecoinformatics' as a term and subfield of ecology first emerged from the biodiversity informatics initiatives of the U.S. Long-term Ecological Research Network (LTER) and the U.S. National Center for Ecological Analysis and Synthesis (NCEAS) in the late 1990s and early 2000s (Kareiva 2001; Brunt et al. 2002). The term was preceded by 'biodiversity informatics', which has been in use since 1992 and gained prominence with Bisby's much-cited article of 2000. Biodiversity informatics and ecoinformatics are largely overlapping fields, the first giving somewhat more emphasis to the taxonomic position of the analysed species, the second more to interactions among taxa and between taxa and with their abiotic environment. Each discipline now has its own journal: *Biodiversity Informatics* started in 2004 and *Ecological Informatics* in 2006.

The International Association for Vegetation Science (IAVS) and its *Journal of Vegetation Science* have been a driving force in the development of ecoinformatics. Ecoinformatics was introduced to IAVS through a lecture on 'Ecoinformatics and the future of community ecology' by R.K. Peet in 2003. IAVS established a Working Group for Ecoinformatics at its 2003 annual meeting, and symposia on the topic were held at its annual meetings in Kona, Hawaii in 2004 and Lisbon in 2005 (see the Special Feature by Bekker et al. 2007). The *Journal of Vegetation Science* has also published some highly-cited descriptions of major tools in ecoinformatics for vegetation scientists, in particular the computer programmes TURBOVEG (Hennekens & Schaminée 2001) for storage of vegetation-plots and JUICE (Tichý 2002) for analysing them.

Contributions in this Special Feature

Ecoinformatics is an essential prerequisite for both retrospective and predictive global change research. In order to detect changes in species composition and vegetation structure, monitoring data have to be collected, archived and distributed in ways that allow relocation of plots, repeated sampling of comparable data and statistical analysis. Generalisation of temporal trends and their spatial patterns relies on large systems of permanent plots and on the joint analysis of numerous time series. Ecoinformatics tools like data exchange standards (Wiser et al. 2011) and networks of data resources (Dengler et al. 2011; Janßen et al. 2011; Lopez-Gonzalez et al. 2011) are indispensable if we are to move from a mere qualitative review of case studies to broad-scale, data-driven meta-analysis.

Research on climate change includes not only efforts to detect its effects, but also predictive modelling in which ecoinformatics provides data on environmental conditions and species distributions. Modelling approaches strongly depend on the availability of recent, accurate and comprehensive data. Assessing the impact of global change boosted the development of adequate models but relevant biological input data are still sparse. Georeferenced point or grid data, however, can be related to a wealth of climatic variables in such models. In doing so, the origin and limitations of model-derived climate data and their resolution relative to fine-scale climatic gradients, as exist in mountainous terrain, have to be borne in mind (e.g. Randin et al. 2009).

Detailed climate scenarios are now easier to obtain than comprehensive species distribution data. This difficulty explains the continued popularity of *Atlas Florae Europaeae* data (http://www.luomus.fi/english/botany/afe/index.htm), despite their low spatial resolution and incomplete coverage of plant families. The current focus of biodiversity informatics is on data in natural history collections, which have been shown to be less effective in providing high-resolution species occurrence data than ecological observations such as those from vegetation-plot data (Garcillán & Ezcurra 2011). Comprehensive, nationwide plot databases give countries like Switzerland, New Zealand, the Netherlands and the Czech Republic a competitive edge in vegetation and global change research. Most parts of the world still lack comprehensive

Journal of Vegetation Science
Doi: 10.1111/j.1654-1103.2011.01313.x

vegetation databases, and supranational initiatives like the Nordic Vegetation Survey (Lawesson et al. 1997) or SynBioSys Europe (Schaminée et al. 2007) are not yet fully developed or sufficiently funded. The fact that development and distribution of comprehensive plot data systems are not yet a major funding priority in most counties suggests that while global change research is often well funded, investments into the capture and delivery of biodiversity data tend to be disproportionately small.

As with species data, the availability of certain environmental variables lags behind that of climate data. For example, soil properties are rarely used in large-scale species distribution modelling (but see Coudon & Gégout 2006; Pompe et al. 2008, 2010), not because they are unimportant, but because they are difficult to obtain at high resolution over large extents. The generation of comprehensive predictor sets and linkage of them to species distribution data are important aspects of ecoinformatics, as demonstrated by Wamelink et al. (2011), Mellert et al. (2011) and Falk & Mellert (2011) in this issue. In turn, vegetation responses may guide the identification of the most meaningful, parsimonious aggregation of predictor variables where we are faced with an overabundance of environmental data (Reger et al. 2011). Correlations between species composition and the environment in databases that comprise both types of information for a large number of plots can be used for novel, advanced modes of interpolation. Ohmann et al. (2011) use species-environment relationships in calibration data to predictively map complex vegetation patterns across large areas. Wamelink et al. (2011) use a similar logic to provide estimates of the abiotic conditions that can be tolerated by community types, thus providing crucial information on the susceptibility of protected habitat types to changing environments (Bittner et al. 2011).

In the Special Feature we have arranged the contributions into five broad topics, starting with an overview of available databases (Dengler et al. 2011), followed by software tools for databasing and data exchange (Janßen et al. 2011; Lopez-Gonzalez et al. 2011; Wiser et al. 2011), species distribution modelling (Falk & Mellert 2011; Mellert et al. 2011; Rupprecht et al. 2011), and modelling and mapping of plant communities (Ohmann et al. 2011; Reger et al. 2011; Rocchini et al. 2011). The final four papers address different issues of the effects of global change on vegetation (Bittner et al. 2011; Czúcz et al. 2011; Voss et al. 2011; Wamelink et al. 2011).

Outlook

Bioinformatics has been the focus of much research activity, funding and publication over the last 15 years, but has also become progressively more equated with applying informatics approaches to molecular data. Although really part of bioinformatics in the true sense of the term (Jones et al. 2006), ecoinformatics has thus far gained only limited visibility. With this Special Feature, we show that the potential of ecoinformatics is much higher than reflected in the current literature. While species occurrence data have received considerable attention and are widely available through large databases such as the Global Biodiversity Information Facility (GBIF; see Wheeler 2004), plot-scale co-occurrence data have received less attention, yet are critically important for understanding the structure of ecological communities and predicting their future change. The amount of information stored in co-occurrence databases is enormous, and with the combined information on species that co-occur at specific locations, site-specific environmental data and structural data, vegetation-plot databases present analytical opportunities that go far beyond those provided by simple occurrence data (Dengler et al. 2011).

To make full use of the available data, further steps are necessary. Here, the development of Veg-X (Wiser et al. 2011) is particularly promising as it provides a data exchange standard that facilitates the exchange of complex vegetation-plot data between databases of different structure and their integration into large data archives such as GBIF, and in addition allows analytical tools and systems to exploit a much broader array of data. Beyond this, we particularly see the following needs for the years ahead:

- More incentives to digitize and share existing data (e.g. Dengler et al. 2011; Janßen et al. 2011).
- Better tools to account for the typical, non-random sampling of vegetation-plot data by resampling the database (e.g. Knollová et al. 2005) or by imputing missing values (e.g. Ohmann et al. 2011), while keeping in mind that there is no a single correct random sampling scheme, but that the adequacy of a distribution of plots strongly depends on the question on hand (Albert et al. 2010).
- An infrastructure for merging datasets where taxon names derive from different sources as applied by different workers at different times. The ambiguity associated with one taxon name being applied to multiple taxon concepts (interpretations of a taxon, as in the sets of specimens that various authors would assign to it), and one taxon concept having multiple possible names, makes data integration complex at best (see Franz et al. 2008; Jansen & Dengler 2010; Patterson et al. 2010), and for large-scale data integration nearly impossible without such an infrastructure.
- A public database that goes beyond the Global Index of Vegetation-Plot Databases (GIVD; Dengler et al. 2011) in providing access to descriptive information about individual plots so that relevant plots can be identified, requested and acquired from appropriate archives.

Already today ecoinformatics can contribute substantially to the global change debate as it allows 'making informed decisions on conservation of biodiversity and sustainable environments' (Recknagel 2011), but by using the tools presented in this Special Feature and addressing the issues mentioned, it has the potential to considerably increase its impact and importance. We can look forward to a time in the near future when ecoinformatics will probably be as central and influential for pure and applied science as the other subdisciplines of bioinformatics.

Acknowledgements

We thank the German Federal Agency of Nature Conservation (BfN), the project BIOTA Southern Africa, and the Floristisch-Soziologische Arbeitsgemeinschaft e. V. (FlorSoz) for their financial support of the meeting in Hamburg where the included articles were presented, and the Chief Editors of the *Journal of Vegetation Science* for making this Special Feature possible.

References

Albert, C.H., Yoccoz, N.G., Edwards, T.C. Jr., Graham, C.H., Zimmermann, N.E. & Thuiller, W. 2010. Sampling in ecology and evolution – bridging the gap between theory and practice. *Ecography* 33: 1028–1037.

Bekker, R.M., Bruelheide, H. & Woods, K. (eds.) 2007. Long-term datasets: From descriptive to predictive data using ecoinformatics. *Journal of Vegetation Science* 18: 457–570.

Bisby, F.A. 2000. The quite revolution: biodiversity informatics and the internet. *Science* 289: 2309–2312.

Bittner, T., Jaeschke, A., Reineking, B. & Beierkuhnlein, C. 2011. Comparing modelling approaches at two levels of biological organisation – Climate change impacts on selected Natura 2000 habitats. *Journal of Vegetation Science* 22: 699–710.

Brunt, J.W., McCartney, P., Baker, K. & Stafford, S.G. 2002. The future of ecoinformatics in long term ecological research. In: Callaos, N. (ed.) *Proceedings of the 6th World Multiconference on Systemics, Cybernetics and Informatics: SCI 2002; July 14–18, 2002 Orlando, Florida, USA. Volume 7: Information systems development, II*. pp. 367–372. International Institute of Informatics and Systemics, Orlando, FL, US.

Canhos, V.P., Souza, S., Giovanni, R. & Canhos, D.A.L. 2004. Global biodiversity informatics: setting the scene for a "new world" of ecological modelling. *Biodiversity Informatics* 1: 1–13.

Coudon, C. & Gégout, J.-C. 2006. Soil nutritional factors improve models of plant species distribution: an illustration with *Acer campestre* (L.) in France. *Journal of Biogeography* 33: 1750–1763.

Czúcz, B., Csecserits, A., Botta-Dukát, Z., Kröel-Dulay, G., Szabó, R., Horváth, F. & Molnár, Z. 2011. An indicator framework for climatic adaptive capacity of natural ecosystems. *Journal of Vegetation Science* 22: 711–725.

Dengler, J., Jansen, F., Glöckler, F., Peet, R.K., De Cáceres, M., Chytrý, M., Ewald, J., Oldeland, J., Finckh, M., Lopez-Gonzalez, G., Mucina, L., Rodwell, J.S., Schaminée, J.H.J. & Spencer, N. 2011. The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. *Journal of Vegetation Science* 22: 582–597.

Dengler, J., Chytrý, M., Ewald, J., Finckh, M., Jansen, F., Lopez-Gonzalez, G., Oldeland, J., Peet, R.K. & Schaminée, J.H.J. (eds.) in press. *Vegetation databases for the 21st century*. Biocentre Klein Flottbek and Botanical Garden, Hamburg, DE.

Ewald, J., Dengler, J. & Finckh, M. 2010. Bericht von der 9. internationalen Tagung zu Vegetationsdatenbanken mit dem Schwerpunkt "Klimawandel" in Hamburg. *Tuexenia* 30: 489–492.

Falk, W. & Mellert, K.H. 2011. Species distribution models as a tool for forest management planning under climate change: risk evaluation of *Abies alba* in Bavaria. *Journal of Vegetation Science* 22: 621–634.

Franz, N.M., Peet, R.K. & Weakley, A.S. 2008. On the use of taxonomic concepts in support of biodiversity research and taxonomy. In: Wheeler, Q.D. (ed.) *The new taxonomy*. pp. 63–86. CRC Press, Boca Raton, FL, US.

Garcillán, P.P. & Ezcurra, E. 2011. Sampling procedures and species estimation: testing the effectiveness of herbarium data against vegetation sampling in an oceanic island. *Journal of Vegetation Science* 22: 273–280.

Guralnick, R. & Hill, A. 2009. Biodiversity informatics: automated approaches for documenting global biodiversity patterns and processes. *Bioinformatics* 25: 421–428.

Hennekens, S.M. & Schaminée, J.H.J. 2001. TURBOVEG, a comprehensive data base management system for vegetation data. *Journal of Vegetation Science* 12: 589–591.

Hintermann, U., Weber, D. & Zangger, A. 2000. Biodiversity monitoring in Switzerland. *Schriftenreihe für Lanschaftspflege und Naturschutz* 62: 47–58.

Janßen, T., Schmidt, M., Dressler, S., Hahn, K., Hien, M., Konaté, S., Lykke, A.M., Mahamane, A., Sambou, B., Sinsin, B., Thiombiano, A., Wittig, R. & Zizka, G. 2011. Addressing data property rights concerns and providing incentives for collaborative data pooling: the West African Vegetation Database. *Journal of Vegetation Science* 22: 614–620.

Jansen, F. & Dengler, J. 2010. Plant names in vegetation databases – a neglected source of bias. *Journal of Vegetation Science* 21: 1179–1186.

Jones, M.B., Schildhauer, M.P., Reichman, O.J. & Bowers, S. 2006. The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution and Systematics* 37: 519–544.

Kareiva, P. 2001. Ecoinformtics: fascilitating access to existing data sets. *Trends in Ecology and Evolution* 16: 226.

Knollová, I., Chytrý, M., Tichý, L. & Hájek, O. 2005. Stratified resampling of phytosociological databases: some strategies

Doi: 10.1111/j.1654-1103.2011.01313.x

for obtaining more representative data sets for classification studies. *Journal of Vegetation Science* 16: 479–486.

Lawesson, J.E., Diekmann, M., Eilertsen, O., Fosaa, A.M. & Heikkilä, H. 1997. The Nordic vegetation survey – concepts and perspectives. *Journal of Vegetation Science* 8: 455–458.

Lopez-Gonzalez, G., Lewis, S.L., Burkitt, M. & Phillips, O. 2011. ForestPlots.net: a web application and research tool to manage and analyse tropical forest plot data. *Journal of Vegetation Science* 22: 610–613.

Mellert, K.H., Fensterer, V., Küchenhoff, H., Reger, B., Kölling, C., Klemmt, H.J. & Ewald, J. 2011. Hypothesis-driven species distribution models for tree species in the Bavarian Alps. *Journal of Vegetation Science* 22: 635–646.

Ohmann, J.L., Gregory, M.J., Henderson, E.B. & Roberts, H.M. 2011. Mapping gradients of community composition with nearest-neighbour imputation: extending plot data for landscape analysis. *Journal of Vegetation Science* 22: 660–676.

Ozinga, W.A., Bekker, R.M., Schaminée, J.H.J. & van Groenendael, J.M. 2004. Dispersal potential in plant communities on environmental condiditions. *Journal of Ecology* 92: 767–777.

Patterson, D.J., Cooper, J., Kirk, P.M., Pyle, R.L. & Remsen D, P. 2010. Names are key to the big new biology. *Trends in Ecology and Evolution* 25: 686–691.

Pompe, S., Badeck, F.-W., Hanspach, J., Klotz, S., Thuiller, W. & Kühn, I. 2008. Projecting impact on plant distributions under climate change–a case study from Germany. *Biology Letters* 4: 564–567.

Pompe, S., Hanspach, J., Badeck, F.-W., Klotz, S., Bruehlheide, H. & Kühn, I. 2010. Investigating habitat-specific plant species pools under climate change. *Basic and Applied Ecology* 11: 603–611.

Randin, C.F., Engler, R., Normand, S., Zappa, M., Zimmermann, N.E., Pearman, P.B., Vittoz, P., Thuiller, W. & Guisan, A. 2009. Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology* 15: 1557–1569.

Recknagel, F. 2011. Ecological informatics: a discipline in the making. *Ecological Informatics* 6: 1–3.

Reger, B., Kölling, C. & Ewald, J. 2011. Modelling effective thermal climate for mountain forests in the Bavarian Alps: which is the best model? *Journal of Vegetation Science* 22: 677–687.

Rocchini, D., McGlinn, D., Ricotta, C., Neteler, M. & Wohlgemuth, T. 2011. Landscape complexity and spatial scale influence the relationship between remotely sensed spectral diversity and survey based plant species richness. *Journal of Vegetation Science* 22: 688–698.

Rupprecht, F., Oldeland, J. & Finckh, M. 2011. Modelling potential distribution of the threatened tree species *Juniperus oxycedrus*: how to evaluate the predictions of different modelling approaches? *Journal of Vegetation Science* 22: 647–659.

Sala, O.E., Chapin, F.S. III, Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., LeRoy Poff, N., Sykes, M.T., Walker, B.H., Walker, M. & Wall, D.H. 2000. Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.

Schalk, P.H. 1998. Management of marine natural resources through by biodiversity informatics. *Marine Policy* 22: 269–280.

Schaminée, J.H., Hennekens, S.M. & Ozinga, W.A. 2007. Use of the ecological information system SynBioSys for the analysis of large datasets. *Journal of Vegetation Science* 18: 463–470.

Soberón, J. & Peterson, T. 2004. Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society B* 359: 689–698.

Tichý, L. 2002. JUICE, software for vegetation classification. *Journal of Vegetation Science* 13: 451–453.

Vitousek, P.M. 1994. Beyond global warming: ecology and global change. *Ecology* 75: 1861–1876.

Voss, N., Simmering, D., Peppler-Lisbach, C., Durka, W. & Eckstein, R.L. 2011. Vegetation databases as a tool to analyze factors affecting the range expansion of the forest understory herb *Ceratocapnos claviculata*. *Journal of Vegetation Science* 22: 726–740.

Wamelink, G.W.W., Goedhart, P.W., Malinowska, A.H., Frissel, J.Y., Wegman, R.J.M., Slim, P.A. & van Dobben, H.F. 2011. Ecological ranges for pH and NO_3 of syntaxa: a new basis for the estimation of critical loads for acid and nitrogen deposition. *Journal of Vegetation Science* 22: 741–749.

Wheeler, Q.D. 2004. What if GBIF? *BioScience* 54: 717–717.

Whitlock, M.C. 2010. Data archiving in ecology and evolution: best practices. *Trends in Ecology and Evolution* 26: 61–65.

Wiser, S.K., Spencer, N., De Cáceres, M., Kleikamp, M., Boyle, B. & Peet, R.K. 2011. Veg-X – An exchange standard for plot-based vegetation databases. *Journal of Vegetation Science* 22: 598–609.