Review Article

Big-data projects: English landscapes and identities

Stephen Rippon[*]

* Department of Archaeology, University of Exeter, UK (✉ s.j.rippon@exeter.ac.uk)

CHRIS GOSDEN, CHRIS GREEN, ANWEN COOPER, MIRANDA CRESWELL, VICTORIA DONNELLY, TYLER FRANCONI, ROGER GLYDE, ZENA KAMASH, SARAH MALLET, LAURA MORLEY, DANIEL STANSBIE & LETTY TEN HARKEL. 2021. *English landscapes and identities: investigating landscape change from 1500 BC to AD 1086*. Oxford: Oxford University Press; 978-0-19-887062-3 hardback £90.

CHRIS GREEN & MIRANDA CRESWELL. 2021. *The shaping of the English landscape: an atlas of archaeology from the Bronze Age to Domesday Book*. Oxford: Archaeopress; 978-1-80-327061-6 Open Access.

This review considers two books outlining the results of a major big-data project in England that sought to make sense of the growing amount of information from developer-funded archaeology, the reporting of finds to the Portable Antiquities Scheme, and from a wide range of other databases. The result is fascinating and thought-provoking discussions of how we could interpret regional variation in archaeological data, although methodological issues present an interesting case study of the challenges that big-data projects face. The publication strategy—of two separate volumes—also raises questions about how we should disseminate the results of large-scale research programmes.

The combined forces of planning guidance, the resulting increase in developer-funded archaeology, and the creation of online databases such as the Portable Antiquities Scheme and Historic Environment Records mean that we now have vast quantities of data, resulting in one of the major challenges facing our profession: how to make sense of it all. In recent years, a series of big-data projects have tried to produce syntheses of particular periods, but one stands out as being more ambitious than the rest: the *English Landscapes and Identities Project* (EngLaID), led by Chris Gosden. The recent publication of two books, *English landscapes and identities* by Chris Gosden, Chris Green and the EngLaID team and *The shaping of the English landscape* by Chris Green and Miranda Creswell, provides the ideal occasion to review the opportunities and challenges that big-data research presents. Together, these two books explore various themes in the development of the English landscape from the Middle Bronze Age to the early medieval period, using a database comprising 900 812 records from 92 existing data repositories, including Historic England's National Record of the Historic Environment, the National Mapping Programme, Excavation Index, Archaeological Investigations Project, county-based Historic Environment Records and the Portable Antiquities Scheme.

English landscapes and identities is a thematic volume that presents an overall discussion of the project's results. Part I contains two papers that introduce the nature of the archaeological data and databases in England and how the EngLaID team went about exploring them. Part II contains five broad thematic chapters that explore "Long Term Interactions Between Society and Ecology", "Movement", "Substances and Cycles" (i.e. food production), "Field systems, Orientation, and Cosmology", and "Identity, Naming, and Division". Part III contains three concluding chapters that consider the core themes of "Scale" and "Time", along with "Conclusions and Reflections".

The thematic chapters in Parts II and III are fascinating and thought provoking. Complex issues, such as the relationship between social agency and environmental determinism, are addressed head on, with discussions of climate change, soil erosion and clearance/land-use history. Some of the analyses arrive at clear, if not unsurprising, results—such as how the construction of the Roman road network transformed mobility across the landscape—while others reveal how little we know, such as the extent to which rivers were navigated (a key conclusion being that just because the Thames was used extensively, it does not necessarily mean that all the others were too). Almost all the chapters combine national overviews with local, in-depth case studies, the exception being the somewhat opaquely titled chapter on "Substances and Cycles" that explores food production in just three small areas in the Thames Basin (the Upper Thames Valley, Middle Thames Valley and the HS1 Corridor in Kent). While it is a pity that these case-study areas were so clustered, as plenty of palaeoenvironmental data exists outside of the Thames Basin, the picture that emerges—of subtly differing farming regimes in different districts—is fascinating.

The chapter on field systems, orientation and cosmology raises the interesting possibility that a significant minority of prehistoric and Roman co-axial field systems could be aligned with the winter solstice. This decision in *English landscapes and identities*—to present chapters containing a brief general discussion that are then followed by detailed case studies of specific locations—does mean that this book has a very different approach to other recently published big-data projects that focus on the national and regional picture. This is not a bad thing, however, as it is important to study landscapes at different scales. This structure is also found in the following two chapters, which consider broad themes that underlie the entire project: "Scale" (both spatial and temporal) and "Time" (the key issues being the definition and understanding of continuity and change, and the significance of place and history to people in the past). The final chapter reflects on the project's overall achievements—that are numerous and impressive—but includes the statement that "Perhaps foolishly, we glossed our attempt at understanding regional variation as 'identity'" (p. 399). It is indeed striking that this important topic is not as prominent as this reader had expected it to be, with just six pages mentioning the hugely important and contentious topic of Romanisation, for example. The opportunity of exploring the concept of *pays*—so well-developed in other landscape studies—was also missed, with only four pages mentioning it.

The second volume produced by the EngLaID team is *The shaping of the English landscape: an atlas of archaeology from the Bronze Age to Domesday Book*. This has a larger format (A4), which allows the maps to be reproduced at a larger scale than in *English landscapes and identities*. The *atlas* has a standard structure throughout, with one or more illustration per page accompanied by a small piece of text describing what is shown. The volume has the feel of a supplement to *English*

landscapes and identities, and this reader could not help feeling that the two books could have been combined into a single A4 volume (which would still have been smaller than those published by other big-data projects, such as the *Rural Settlement of Roman Britain*).

The illustrations vary in quality across both volumes. In many cases, the colour scheme is very poor, with small, mid-red dots being difficult to see against a mid-grey background, whereas others use strikingly contrasting colours and so are far easier to read. A particular issue in *English landscapes and identities* is that various illustrations are too small and cluttered: Figure 7.3, for example, has 40 maps of England crammed on to a print area that is less than A5 in size, resulting in maps that are simply illegible.

Many of the maps, in both volumes, do convey fascinating insights into the landscape and will provide a resource of lasting value, while others may be of less use. The map of "Fortifications – Roman" (*atlas*, p. 74: there are no figure numbers in this volume), for example, contains a mass of undifferentiated red dots that appears to include everything from short-lived marching camps from the Conquest period, forts occupied for a few decades in the mid- to late first century AD, to Late Roman forts of the Saxon Shore. I cannot see the use for such a map that contains sites of such different character and from completely different phases of Roman Britain. The maps (in both volumes) must also be read with very great care, due to what is the most fundamental issue underlying this project: how it appears to have simply accepted other people's interpretations of archaeological data without any critical analysis of whether those interpretations are correct. The map of "Settlement – Roman" (*atlas*, p. 30), for example, contains a series of 'villas' in Devon for which there is no evidence whatsoever, other than a few fragments of ceramic roof tile, some of which are re-used in medieval buildings. Are we really happy to say such sites are villas? The problem is that the EngLaID team did not 'clean' the data (see *English landscapes and identities*, pp. 33–34), but simply accepted past interpretations of evidence in archives such as Historic Environment Records without critically assessing the actual primary data. The result is that—to this reader at least—maps such as the "Settlement – Roman" are seriously misleading.

The success of developer-funded archaeology has meant that, as a profession, archaeologists must start making sense of a growing amount of data. Other 'big-data' projects have chosen to limit their scope and focus on specific periods, or on the transition from one period to another, and, as such, EngLaID is by far the most ambitious. But herein lies what for this reader is the biggest question: did EngLaID gather too much data, with the result that there was insufficient time to clean it all (e.g. deleting 'villas' for which the evidence is just a few pieces of Roman tile re-used in medieval buildings), and, in so doing, reduce the integrity of the maps? The decision to publish two separate volumes also seems strange, as the reader who consults only the *atlas* will not have a clear understanding of the methodological issues and contextual discussion provided in *English landscapes and identities*. Indeed, the *atlas* does not even appear in the Bibliography of *English landscapes and identities*, and the latter contains very few references to the former. A more joined-up approach towards publication would have been better. These nagging worries about how the project has been published, however, must be balanced against the very important contribution that it will make to our understanding of landscape through its general overviews and detailed case studies. The chapters all provide fascinating discussions, and the undoubtedly difficult decisions regarding methodology, structure and publication strategy are themselves food for thought.