



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/155752/>

Version: Accepted Version

---

**Article:**

Watkins, J.H.R., Cameron, R.W.F., Sjöman, H. et al. (2020) Using big data to improve ecotype matching for Magnolias in urban forestry. *Urban Forestry & Urban Greening*, 48. 126580. ISSN: 1618-8667

<https://doi.org/10.1016/j.ufug.2019.126580>

---

Article available under the terms of the CC-BY-NC-ND licence  
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# 1 Using big data to improve ecotype matching for Magnolias in 2 urban forestry

## 3 Summary

4 Trees play major roles in many aspects of urban life, supporting ecosystems, regulating temperature  
5 and soil hydrology, and even affecting human health. At the scale of the urban forest, the qualities of  
6 these individual trees become powerful tools for mitigating the effects of, and adapting to climate  
7 change and for this reason attempts to select the right tree for the right place has been a long-term  
8 research field. To date, most urban forestry practitioners rely upon specialist horticultural texts (the  
9 heuristic literature) to inform species selection whilst the majority of research is grounded in trait-  
10 based investigations into plant physiology (the experimental literature). However, both of these  
11 literature types have shortcomings: the experimental literature only addresses a small proportion of  
12 the plants that practitioners might be interested in whilst the data in the heuristic (obtained through  
13 practice) literature tends to be either too general or inconsistent. To overcome these problems we used  
14 big datasets of species distribution and climate (which we term the observational literature) in a case  
15 study genus to examine the climatic niches that species occupy in their natural range. We found that  
16 contrary to reports in the heuristic literature, *Magnolia* species vary significantly in their climatic  
17 adaptations, occupying specific niches that are constrained by trade-offs between water availability  
18 and energy. The results show that not only is ecotype matching between naturally-distributed  
19 populations and urban environments possible but that it may be more powerful and faster than  
20 traditional research. We anticipate that our findings could be used to rapidly screen the world's woody  
21 flora and rapidly communicate evidence to nurseries and plant specifiers. Furthermore this research  
22 improves the potential for urban forests to contribute to global environmental challenges such as  
23 species migration and ex-situ conservation.

## 24 **Keywords**

25 Big data; Biogeography; Ecotype matching; Predictive ecology; Urban trees.

## 26 **Introduction**

27 Through their provision of a complex suite of ecosystem services such as run-off water management,  
28 biodiversity habitat and cultural services, urban forests play a key role in mitigating some of the  
29 effects of a changing climate (Ordóñez Barona, 2015; Wilson, 2016; Acuto *et al.*, 2018). However, if  
30 urban forests are to be able to provide these benefits, their fitness needs to be improved so that they  
31 are able to deal with the many stresses that reduce urban forest growth and increase mortality risk  
32 (Bialecki, Fahey and Scharenbroch, 2018) and that are being exacerbated under climate change, such  
33 as prolonged or aseasonal drought, flooding or pathogens (Roloff, Korn and Ñ, 2009; Allen,  
34 Breshears and McDowell, 2015; Fuller and Quine, 2016).

35

36 Building on earlier discourse (Santamour, 1990), urban forestry researchers and practitioners have  
37 emphasised the importance of species selection and diversification as a means to achieve this (Krajter  
38 Ostoić and Konijnendijk van den Bosch, 2015; Morgenroth *et al.*, 2016) with a number of recent  
39 publications providing guidance to aid decision making at the practice level (Vogt *et al.*, 2017;  
40 Barbrook *et al.*, 2018; Hirons and Sjoman, 2018). Whilst these publications are a significant  
41 development, they inevitably have to compromise between the detail with which they can present  
42 information and the range of species they are able to discuss, compounded by the practical limits of  
43 what genetic material nurseries have access to. Exacerbating these constraints, economic pressures on  
44 horticultural production lead to increasingly reduced genetic diversity amongst the trees available in  
45 nurseries, with many species represented by either a single clone or a small number of seed orchards.  
46 Some exceptions to this exist for species such as *Acer rubrum*, where there are multiple named clones,  
47 but many selections are based primarily on aesthetic criteria such as autumn colour potential and habit  
48 rather than fitness to environment. If specifiers want to truly diversify the gene pool of urban forests,  
49 new tools are required that can identify urban-fit ecotypes.

50

### 51 *Criteria for a case study genus*

52 This study uses a single genus as a case study to investigate the literature for the efficacy of selecting  
53 novel ecotypes, firstly reviewing existing species selection literature, then developing a methodology  
54 for assessing new sources of information. The criteria for a suitable candidate genus should reflect the  
55 nature of the challenge: it should not be widely planted in urban forests, nor widely discussed in urban  
56 forestry literature, and ideally it should display some degree of natural variation. *Magnolia* is a  
57 flagship genus (Cires *et al.*, 2013) that fits these criteria, with high ornamental value and perceived  
58 low tolerance of stress or disturbance. The literature of *Magnolia* tends to focus on horticultural or  
59 cultural aspects (Bunting 2016, Callaway 1994, Gardiner 2000), genetics and phylogeny (Muranishi  
60 *et al.*, 2013; Budd, Zimmer and Freeland, 2015), and conservation (Cicuzza, Newton and Oldfield,  
61 2007; Rivers *et al.*, 2016), with relatively few studies into their functional traits (Cires *et al.*, 2013) or  
62 the extent to which species within the genus are able to withstand stresses (Sjöman, Hirons and  
63 Bassuk, 2018). Nevertheless, *Magnolia* might contain effective selections for urban forestry given its  
64 reported intra-specific variation (Azuma, Toyota and Asakawa, 2001; Azuma *et al.*, 2011) and the  
65 wide range of environmental conditions to which it has adapted (Azuma *et al.*, 2001). In spite of its  
66 reputation for being intolerant of climatic extremes and poor quality soils, its ornamental qualities are  
67 highly valued and could be important in encouraging people to accept urban forests as an acceptable  
68 landscape type for dense cities (Hitchmough and Bonugli, 1997; Hoyle, Hitchmough and Jorgensen,  
69 2017).

70

### 71 *The existing species selection literature*

72 A preliminary review identified two broad categories of literature that could be used in species  
73 selection: the experimental and heuristic literature (see Table 1). The experimental literature is rooted  
74 in functional ecology, tends to be published in peer-reviewed journals, and typically studies either  
75 morphological or physiological traits in controlled studies. By contrast, the heuristic literature (by  
76 which we mean work that is based upon the accumulated knowledge of those working in practice in  
77 horticulture or urban forestry (Ippoliti, 2015; Vogel and Henstra, 2015)) describes experiences of  
78 growing a wide range of species and observing their characteristics over a long period of time. The

79 heuristic literature tends to be published in the form of horticultural monographs, nursery catalogues  
80 or growers' manuals and provides information about whole-plant characteristics such as overall size  
81 or growing conditions, or particular ornamental qualities such as leaf or flowering characteristics.  
82 Both literature types have specific objectives and are aimed at different audiences: the experimental  
83 literature, for example, uses technical language, is highly focussed in its study area and is mostly used  
84 by researchers to address macro ecological questions; on the other hand, the heuristic literature  
85 presents a wide range of information that is aimed at the horticultural and professional landscape  
86 sectors. The heuristic literature differs philosophically from the experimental in that observations are  
87 made without the capacity to know what the responses would be if a different set of conditions or  
88 treatments were involved. In practice, urban foresters tend to rely upon a range of sources, with some  
89 publications such as professional journals or industry-endorsed guidance (Hirons and Sjöman, 2018)  
90 straddling the boundaries of these broad categories.

91 *Table 1: A typological classification of the existing urban forestry species selection literature*

<b>Literature typology</b>	<b>Publications</b>	<b>Target audience</b>	<b>Data</b>
Experimental	Trait literature <sup>1</sup>	Functional ecologists, dendrologists, botanists	Functional traits (e.g. SLA, SSD, Plant height) or functional type
Heuristic	Nursery catalogues <sup>2</sup>	Gardeners, landscape architects, landscape contractors, urban foresters	Plant size, floral or leaf aesthetics, resource requirement (eg water, light), soil conditions
Heuristic	Encyclopaedia <sup>3</sup>	Gardeners, landscape architects, urban foresters	Plant size, floral or leaf aesthetics, resource requirement (eg water, light), soil conditions
Heuristic	Horticultural monographs <sup>4</sup>	Gardeners, landscape architects, botanic gardens and arboretums	Plant size, floral or leaf aesthetics, resource requirement (eg water, light), soil conditions
Heuristic	Industry guidance <sup>5</sup>	Landscape architects, landscape contractors, urban foresters	Plant size, floral or leaf aesthetics, resource requirement (eg water, light), soil conditions, management requirements

92

<sup>1</sup> For example, (Kattge *et al.*, 2011; Sjöman, Hirons and Bassuk, 2018).

<sup>2</sup> For example, Glover 2016, or catalogues from UK nurseries such Burncoose and Coblands.

<sup>3</sup> For example, Hillier Manual of Trees and Shrubs (8<sup>th</sup> ed), Dirr (2011) or Gardiner (2012).

<sup>4</sup> For example, Bunting (2016), Callaway (1994), Gardiner (2000), Treseder (1978).

<sup>5</sup> For example, Samson et al (2017) Hirons and Sjöman (2018), UK National Plant Specification.

93 *Alternative data sources: the observational literature*

94 In contrast to the experimental and heuristic literatures which describe plant performance or traits, a  
 95 third literature source exists that could be used by urban foresters to understand the naturally  
 96 occurring niches that plants occupy. Using the type of data used in biogeographical studies (Table 2)  
 97 would require a fundamentally different approach to species selection, requiring urban foresters to  
 98 understand and harness evolutionary adaptations, target specific populations or ecotypes and then  
 99 match these to specific designed environments. Such an approach would enable a far greater degree of  
 100 precision and confidence in designing urban forests to meet specific challenges.

101 *Table 2: Proposed additions to urban forestry species selection literature*

<b>Literature typology</b>	<b>Publications</b>	<b>Target audience</b>	<b>Data</b>
Observational	Plant identification and distribution resources <sup>6</sup>	Taxonomists, conservationists and horticulturists	Natural distribution of species or individuals, habitat in fundamental or realised niche
Observational	Climate <sup>7</sup>	Climate scientists, biogeographers, ecologists, planners	Mean monthly rainfall, mean monthly temperature

102  
 103 We use the term ‘observational literature’ to describe the vast records of observations of plant  
 104 occurrences and climate set out in Table 2. The observational literature category includes all records  
 105 of the natural distribution of species, whether the results of fieldwork, plant collecting or exploration  
 106 and is usually held in herbaria or databases (such as GBIF), whilst climate records can be accessed  
 107 through resources such as WorldClim. Comprising millions of data points, this information is often  
 108 termed ‘big data’ (Hallgren *et al.*, 2016; Serra-Diaz *et al.*, 2017; Allen *et al.*, 2018; Pelletier *et al.*,  
 109 2018), and is increasingly used as a powerful resource for describing species distribution and  
 110 environmental adaptation (Booth, 2018; Wang *et al.*, 2018).

111

---

<sup>6</sup> For example, GBIF (<https://www.gbif.org>), Global Plants (<https://plants.jstor.org>)

<sup>7</sup> For example, Global Climate Data (<http://www.worldclim.org>) or The World Bank Climate Change Knowledge Portal (<http://sdwebx.worldbank.org/climateportal/>)

112 These records are not without their idiosyncrasies: records have been accumulated over the past four  
113 hundred years and whilst these records have often been reviewed regularly by botanists working in  
114 herbaria, they can reveal bias or patchiness in their coverage of a species distribution, level of detail,  
115 or nomenclature. As such, these records are often difficult to interpret, contextualise or physically  
116 access, requiring archival research in herbaria and whilst major efforts are being taken to digitise  
117 these records and share via online repositories, a large proportion of the world's 380m herbarium  
118 vouchers remain un-digitised (James *et al.*, 2018). Similar factors affect climatic data (particularly  
119 rainfall and temperature) that have been recorded around the world over the past 150 years. As a  
120 result, despite the sophisticated interpolation of climate data and rapidly evolving techniques for  
121 recording information, models do not yet offer a consistently accurate record of climate across  
122 multiple scales of resolution, posing problems for identifying climate niches in mountainous areas  
123 where aspect and elevation complicate interpolation.

124

125 In spite of these shortcomings, the theoretical basis for bringing observational literature sources  
126 together is robust: the effects of water and energy relations upon plant distributions has been well  
127 established through indices of potential evapotranspiration, moisture indices and warmth index (Yim  
128 and Kira, 1975; Kreft and Jetz, 2007; Wright *et al.*, 2017), and as such the biological and climate data  
129 that is available online remains a substantial resource. If we are able to treat these resources in a  
130 probabilistic manner using biogeographical conceptual frameworks and techniques, it should be  
131 possible to identify not only variation in bioclimatic niches across which *Magnolia* is distributed but  
132 also population-level intra-specific variation, and thus providing the basis for improved matching  
133 between ecotype and urban site into which it might be planted.

134

135 In this paper we identify a new literature source and develop a methodology for handling the  
136 enormous and widely distributed data sets that it contains, allowing us to address three long-standing  
137 challenges in the management strand of urban forestry (Morgenroth *et al.*, 2016): what is the most  
138 effective source of information for species selection? Is it possible to access information about  
139 superior trees at the level of the ecotype, rather than species? And finally, if these literature sources

140 are sufficiently powerful to identify likely superior ecotypes, how accessible are they to urban  
141 foresters? Together, these research questions allow us to rapidly screen genetic diversity within  
142 species to identify sub-specific populations suitable for urban forestry under climate change.

## 143 **Materials and Methods**

144 To address these challenges, we developed a novel research approach involving a sequence of steps to  
145 classify and analyse two classic literature sources and a new source of species selection literature. To  
146 answer the three research questions identified above, we carried out the following steps:

- 147 a) We described the scope of each literature type, recording the number of species discussed and  
148 the number of records for each species within each source,
- 149 b) We assessed the level of precision to which traits, resource use or climate niche were  
150 described (i.e. genus, inter-specific, or intra-specific), and
- 151 c) We assessed the efficacy of each literature type in identifying potential match between  
152 resource requirement, traits or climate niche and possible designed urban sites.

153 .

### 154 *Identifying sources for each literature typology*

155 Urban forestry literature is highly diverse, with specification sources and practices varying widely  
156 between practitioners. A preliminary literature review was carried out, identifying three broad sources  
157 of literature: the experimental literature, heuristic literature, and observational literature (Tables 1 and  
158 2). Literature searches were tailored for each literature type. For the experimental literature, searches  
159 were carried out on Scopus, Web of Science and Google Scholar using terms including ‘*Magnolia*’,  
160 species names (eg ‘*acuminata*’, ‘*biondii*’, ‘*campbellii*’ etc.), plant organs (e.g., ‘leaf’, ‘stem’, ‘root’)  
161 and traits, including spelling variations and abbreviations (e.g. ‘SLA’ and ‘Specific Leaf Area’, ‘SSD’  
162 and ‘Specific Stem Density’ / ‘Wood Density’), complemented by searches in trait databases (TRY,  
163 Bien R). *Magnolia* species were searched for in the heuristic literature in 12 texts that are frequently  
164 used by landscape architects and urban foresters to account for the varying approaches that urban  
165 foresters take to species selection and their own interests or specialisms. Some well-established  
166 sources of heuristic literature were not eligible for this study due to opaque evaluation or inconsistent  
167 data collection techniques (e.g. the Royal Horticultural Society’s AGM scheme). Within the  
168 observational literature, climate data was searched using the University of East Anglia’s world  
169 climate model (accessed at <http://sdwebx.worldbank.org/climateportal/>), whilst plant records were

170 searched using the online repositories GBIF, Global Plants and the Chinese Online Herbarium, and  
171 supplemented by archival searches in herbaria at Oxford University, the Royal Botanic Gardens  
172 Edinburgh and Kew (RBGE, RBGK). During these searches we followed the convention established  
173 by the IUCN Red List (Rivers *et al.*, 2016) of lumping subspecies and varieties into species accounts  
174 as a means of standardising the analysis across different literature sources, with the exception of *M.*  
175 *sieboldii* where practitioners habitually maintain the distinction between *M. sieboldii* and its  
176 subspecies *M. sieboldii* subsp. *sinensis*.

177

### 178 *Gathering and tabulating data*

179 Within the heuristic literature, information relating to requirements of water and light were considered  
180 more precise than descriptions of hardiness as hardiness is often context-dependent and thus difficult  
181 to interpret consistently. These descriptions were recorded as categorical variables followed by a  
182 review of the vocabulary used in the publication so that numeric values on a scale of 1-5 could be  
183 applied to the categorical variables for resource requirement (1 = low resource requirement, 5 = high  
184 resource requirement), similar to the systems used by Ellenberg (1974) or (Bassuk *et al.*, 2009). For  
185 example, Hillier Manual of Trees and Shrubs (Armitage, Edwards, & Lancaster (eds), 2014) uses the  
186 terms “Good in dry soils,” “Well-drained,” “Moist,” “Plenty of moisture,” and “Wet” to describe  
187 optimal growing conditions. These terms were recorded, ordered, and assigned numeric values to  
188 reflect this order; in this way, “Good in dry soils” was assigned ‘1’ and “Wet” was assigned ‘5’.

189

190 Within the experimental data, the well-established plant economics spectrum identifies key traits that  
191 explain plant metabolism and tolerance of stress, such as specific leaf area and plant height (Wright *et*  
192 *al.*, 2004; Díaz *et al.*, 2016). Data for key traits that play recognised roles in plant functioning (Pierce  
193 *et al.*, 2017) were gathered and recorded, and filtered for data that recorded growth under normal or  
194 control conditions (i.e. data from experimental studies where variables such as drought or soil salinity  
195 were studied were excluded). Data was then formatted to SI units to allow comparison.

196

197 Within the observational literature, only records of verified observations were included in the study,  
198 as reports that describe distribution ranges (such as “between 1800m – 2400m in Sichuan, Henan and  
199 Hubei”) were considered too vague for inclusion. After positively identifying a plant record, the  
200 location of the observation was recorded using Google Maps and decimal coordinates were derived.  
201 The decimal coordinates were then used to identify the location with the University of East Anglia’s  
202 climate model (<http://sdwebx.worldbank.org/climateportal/>) and mean monthly rainfall and  
203 temperature were recorded. Whilst Potential Evapotranspiration (PET) is commonly used in ecotype  
204 matching and biogeographical modelling (Haxeltine and Prentice, 1996), we considered that plotting  
205 water-energy relations found at each site (sensu Aguilar-Romero et al. (2017)) would allow for a  
206 study design that was more sensitive to the relatively small number of population records and  
207 potentially have greater explanatory power than a single variable that integrated both water and  
208 temperature. To represent energy relations, Warmth Index was considered more sensitive than Mean  
209 Annual Temperature (Woodward, Lomas and Kelly, 2004) as this variable accounts for the intensity  
210 of energy during growing season, and excludes cold season temperatures which can have a distorting  
211 effect on an annual mean. On this basis, we summed the mean monthly rainfall to calculate the annual  
212 rainfall and used the mean monthly temperature to calculate Warmth Index using the formula  
213 developed by Yim and Kira (1975):

$$WI = \sum (Tm-5), \text{ when } Tm > 5^{\circ}C$$

(*Tm*: Monthly Mean Temperature)

216

### 217 *Data analysis*

218 Microsoft Excel (v15.26) was used to tabulate the data and RStudio (v1.1.383) was used to  
219 manipulate data and carry out statistical analysis. To calculate the number of species discussed in each  
220 literature type, records were tabulated, ranking species from high to low (see Table 3) within each  
221 literature type (experimental, heuristic and observational). At this point it became clear that there  
222 were insufficient experimental data to identify interspecific differences and this literature source was  
223 excluded from further studies.

224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251

A key cultural perception is that *Magnolia* are generally intolerant of climatic extremes and poor-quality soils, with little variation reported within the genus (Samson *et al.*, 2017). To investigate this, the second step tested the data in the heuristic and observational literatures for normality using the Shapiro-Wilks W test ( $H_0$  = sample distribution is not different from normal distribution), and calculating means for each species' reported resource requirement or the availability of resources in their natural distribution.

Following the results of the Shapiro-Wilks W test, non-parametric analysis of variance was carried out using the Kruskal-Wallis rank sum test to determine whether the means for each species resource requirement or resource availability were significantly different, allowing us to assess the level of inter-specific variation reported in each literature. Due to their different objectives, the heuristic and observational literature reported plant water-energy balances in subtly different ways, resulting in different analyses: in the heuristic literature, preferred provision of water and light were plotted against each other (Figure 5) and using the observational literature, Annual Rainfall (mm) was plotted against Warmth Index (WI) in Figure 6, in effect creating basic Species Distribution Models (SDM). Regression lines and 95% confidence intervals were plotted and the degree of inter-specific variation in each literature type was recorded as  $R^2$ , slope and intercept. At this point it became clear that the heuristic literature did not identify significant variation at the inter-specific level and was excluded from further analysis. The last step in answering our second question (the level of precision that the literature describes genetic variation) was to assess whether the observational literature was capable of identifying intra-specific variation: the same process was repeated for each species, plotting Annual Rainfall against Warmth Index and recording  $R^2$ , slope and intercept of the regression.

To answer the third question (whether the literature might be able to describe fit between an ecotype and an urban environment under climate change), hypotheses for rear and leading edge populations (Hampe and Petit, 2005) were identified using the SDMs for *Magnolia* species that showed significant regression, and these populations were plotted against selected urban environments in

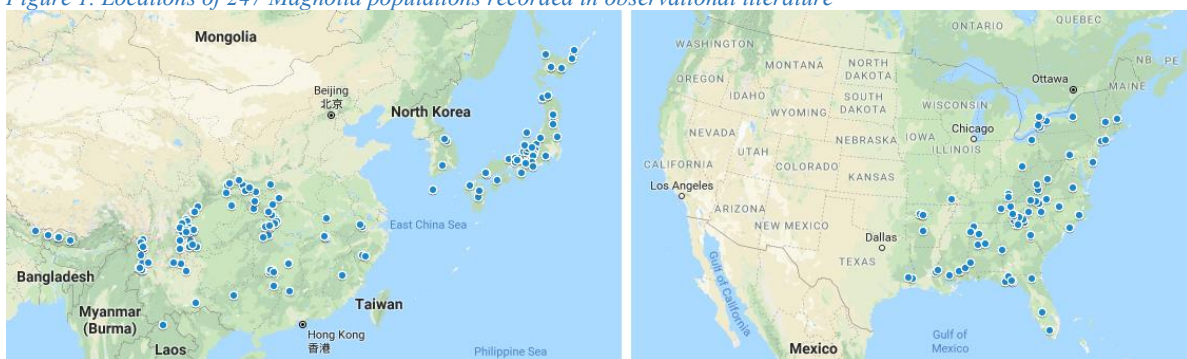
252 Northern Europe to identify potential matches between naturally distributed *Magnolia* populations  
253 and current urban climates.

254 **Results**

255 *Range of species recorded in each literature type*

256 Significant differences between the scope of each literature were identified and are summarised in  
257 Table 3: the heuristic literature discussed most of the temperate *Magnolia* species, suggesting a  
258 thorough treatment of the genus. Within this literature, most of the sources tended to provide generic  
259 descriptions of species' preferred growing conditions and detailed information about horticultural  
260 qualities such as flower size or colour. On the other hand, the experimental literature was relatively  
261 narrow in terms of the species discussed and uneven in the level of detail to which they were  
262 discussed: the large majority of trait data were calculated in controlled studies in north American  
263 universities, with especially high numbers of replicates in the studies of *M. fraseri*, *M. grandiflora*  
264 and *M. virginiana*.

265 *Figure 1. Locations of 247 Magnolia populations recorded in observational literature*



267

268 The observational literature was the most extensive both in terms of species discussed but also in  
269 terms of the level of detail provided, with 247 records identified (Fig 1). It was found that the  
270 observational literature was limited for some taxa, perhaps due to their limited species distribution  
271 (e.g. *M. dawsoniana*), recording bias or geopolitical factors that might affect botanical exploration.  
272 Nevertheless, the studies reported represent a small fraction of the records available within herbaria,  
273 suggesting that it might be possible to develop a stronger and more robust database of occurrences.

274

275 *Table 3. Number of times a species is reported in each literature type*

276

Species	Heuristic literature	Experimental literature	Observational literature
<i>M. acuminata</i> (L.) L	11	5	18
<i>M. biondii</i> Pamp.	6	-	14
<i>M. campbellii</i> Hook. F. & Thomson	11	-	7
<i>M. dawsoniana</i> Rehder & E. H. Wilson	7	-	9
<i>M. denudata</i> Desr,	11	3	9
<i>M. fraseri</i> Walter	-	6	17
<i>M. globosa</i> Hook. F. & Thomson	-	-	8
<i>M. grandiflora</i> L.	13	11	7
<i>M. kobus</i> DC	12	2	16
<i>M. liliiflora</i> Desr.	11	1	6
<i>M. macrophylla</i> Michx.	17	2	16
<i>M. obovata</i> Thunb.	7	4	9
<i>M. officinalis</i> Rehder & E. H. Wilson	7	1	12
<i>M. rostrata</i> W.W.Sm.	4	-	11
<i>M. salicifolia</i> (Siebold & Zucc.) Maxim.	8	2	12
<i>M. sargentiana</i> Rehder & E. H. Wilson	8	-	15
<i>M. sieboldii</i> K. Kobch	18	2	15
<i>M. sprengeri</i> Pamp.	7	-	13
<i>M. stellata</i> (Siebold & Zucc.) Maxim.	12	2	6
<i>M. tripetala</i> (L.) L.	9	4	12
<i>M. virginiana</i> L.	19	6	12
<i>M. wilsonii</i> (Finet & Gagnep.) Rehder	10	-	10

277

278 *Differences between data formats in each literature*

279 Table 4 illustrates the challenges of using the heuristic literature, with criteria for plant behaviour  
280 varying greatly between (or sometimes within) sources. This is complicated by the literary style or  
281 vocabulary that the sources use, often giving the impression of ‘Cinderella’ species that require  
282 difficult-to-achieve conditions of fertile, moist, well-drained soils, and making it difficult to

283 consistently compare identify distinctive features or nuances of species between sources.  
 284 Nevertheless, the use of the 1-5 scale allowed this vocabulary to be compared effectively across  
 285 sources and analysed in later stages.

286 *Table 4. Resource requirements of the eight most commonly-described Magnolia species in the heuristic literature*

Species name	Description
<i>M. acuminata</i>	“Responds to rich living, good drainage and plenty of moisture” Hillier Manual of Trees and Shrubs (8 <sup>th</sup> ed)
<i>M. campbellii</i>	“Quite happy in full sun where moisture and humidity levels are high” Gardiner (2000)
<i>M. denudata</i>	“Prefers moist soils” Callaway (1994)
<i>M. grandiflora</i>	“Needs well-drained soils” Bunting (2016) “Does not like dry soils” Gardiner (2000) “Needs fertile, moist, well-drained soil” Burncoose (2018)
<i>M. kobus</i>	“Adaptable to many conditions” Callaway (1994)
<i>M. liliiflora</i>	“Prefers well-drained soils” Brickell (2003)
<i>M. stellata</i>	“Tolerates shade although it is more vigorous and blooms more profusely in sunny locations” Callway (1994)
<i>M. virginiana</i> var. <i>australis</i>	“Needs medium to wet soils” Missouri Botanic Garden

287  
 288 The values found in the experimental literature (Table 5) demonstrate that whilst these studies discuss  
 289 aspects of plant morphology or physiology that are essential for plant functioning and explain aspects  
 290 of stress tolerance or competitive ability, there is not yet sufficient data to generate meaningful  
 291 findings to guide urban forestry species selection or to ordinate species within functional schemes in  
 292 the manner of Reich (2014) or Grime and Pierce (2012). For this reason, the experimental literature  
 293 was not evaluated further in this study.

294 *Table 5. Mean trait values reported within the experimental literature*

Taxa	Plant height (m)	Leaf area (mm)	Leaf dry matter content (%)	Specific leaf area (mm <sup>2</sup> /mg <sup>-1</sup> )	Specific Stem Density (g/cm <sup>3</sup> )	Leaf turgor loss point ( $\Delta\Psi_{\pi 100}$ (MPa))
<i>M. acuminata</i>	27.6	-	-	-	-	0.40
<i>M. denudata</i>	29.3	-	17.32	29.27	0.43	-
<i>M. fraseri</i>	21.5	-	-	22.70	0.40	-
<i>M. grandiflora</i>	14.1	9185	-	9.72	0.44	-
<i>M. kobus</i>	17.8	-	-	-	-	0.26
<i>M. liliiflora</i>	4.0	-	-	-	-	-
<i>M. macrophylla</i>	23.5	-	-	-	-	-
<i>M. obovata</i>	27.7	-	-	12.38	-	-
<i>M. officinalis</i>	20.0	-	-	-	-	-

<i>M. salicifolia</i>	-	-	-	32.82	-	0.35
<i>M. sieboldii</i>	8.0	-	-	-	-	0.39
<i>M. stellata</i>	6.2	-	-	-	-	-
<i>M. tripetala</i>	13.9	-	-	-	-	0.47
<i>M. virginiana</i>	10.0	6912	-	10.28	0.42	-

295

296

Whilst the heuristic literature describes the preferred or acceptable conditions for plant growth in

297

horticultural environments, the observational literature reports the actual conditions experienced by

298

trees in their natural ranges, showing that there are both greater inter-specific and intra-specific

299

differences in the Warmth Index than the Annual Rainfall experienced by *Magnolia* populations. Figs

300

2a & b illustrate these differences (including London as a benchmark for comparison), showing that

301

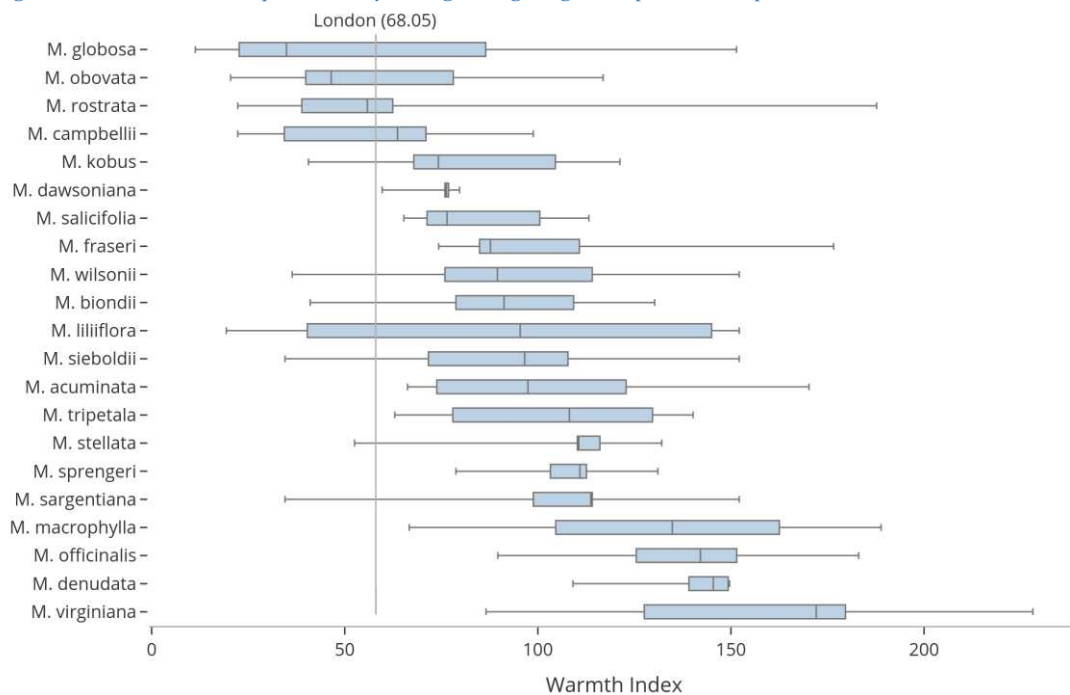
most *Magnolia* populations are likely to grow in conditions that are slightly warmer and generally

302

with much higher water availability than European urban environments,

303

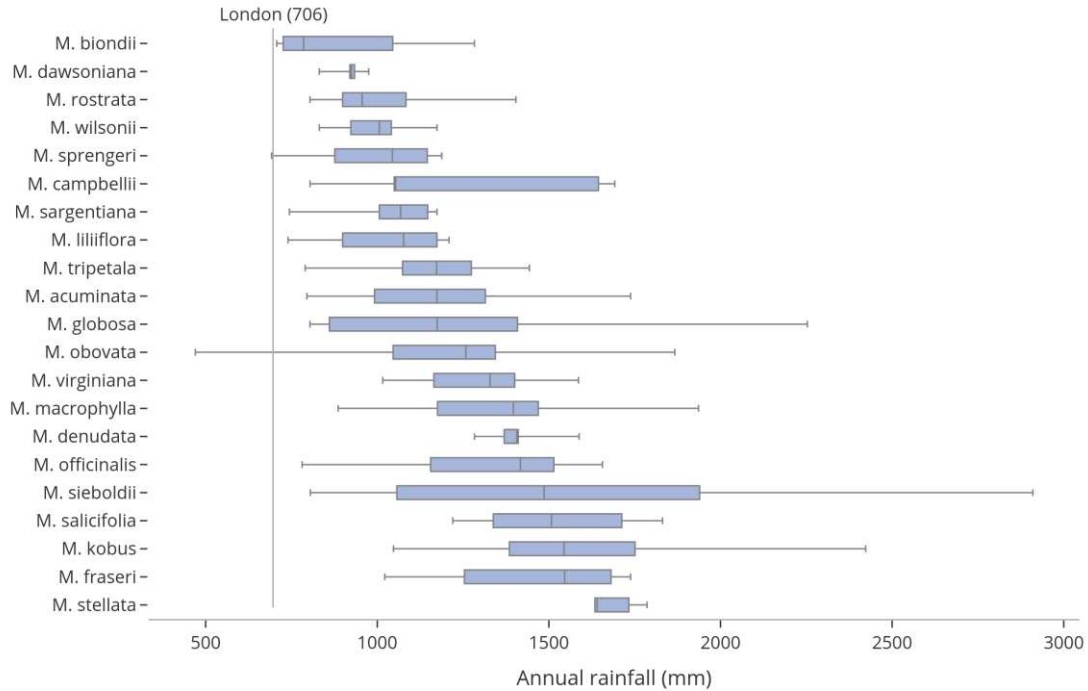
*Figure 2a. Warmth Index experienced by wild-growing Magnolia species, as reported in the observational literature*



304

305

306 *Figure 2b. Annual Rainfall experienced by wild-growing Magnolia species, as reported in the observational literature*



307

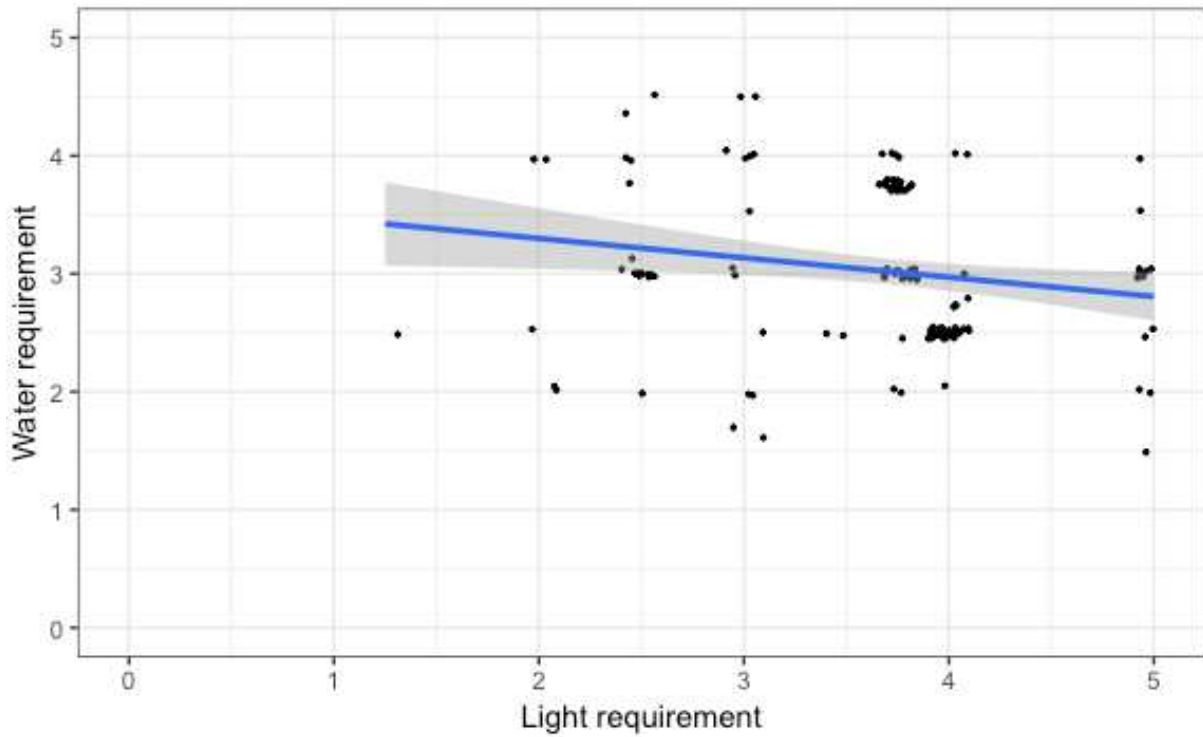
308

309 *Identifying inter- and intra-specific variation*

310 Figure 3 describes the preferred growing conditions for horticultural situations as reported in the  
 311 heuristic literature, demonstrating that although this literature describes a wide range of species, it  
 312 identifies weak inter-specific variation in *Magnolia*, suggesting that most Magnolias are fairly similar  
 313 in a functional sense. Most records suggest that the preferred conditions are for relatively high levels  
 314 of light and water, with little acknowledgement of how these levels might vary in a global context and  
 315 little capacity to identify the limits of stress tolerance that they could endure. By contrast, Figure 4  
 316 identifies not only a range in experienced conditions but also a potential trade-off in the water-energy  
 317 balance.

318  
319

Figure 3. Genus-level description of the preferred conditions for cultivating Magnolias, reported within the heuristic literature

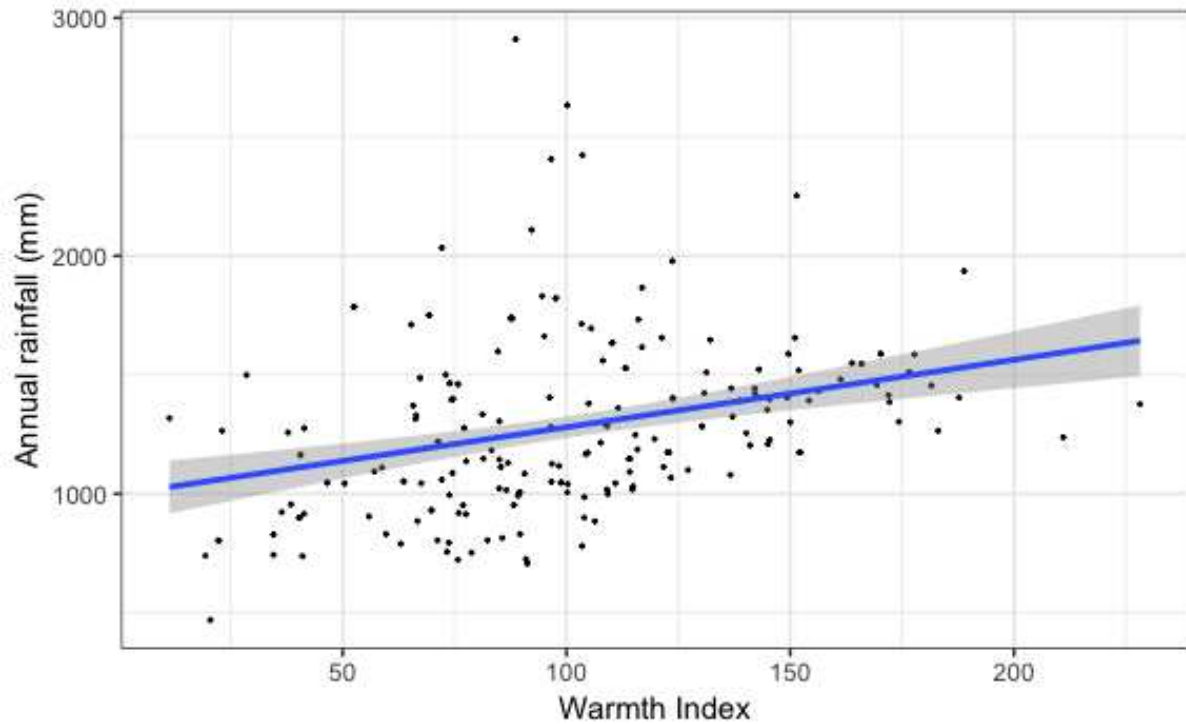


320

321  $r = -0.1968675$ ,  $p \leq 0.1$ ,  $n = 204$ ,  $R^2 = 0.3199$

322

323 Figure 4. Genus-level distribution model for Magnolias in terms of annual rainfall and warmth index, as reported within the  
324 observational literature



325

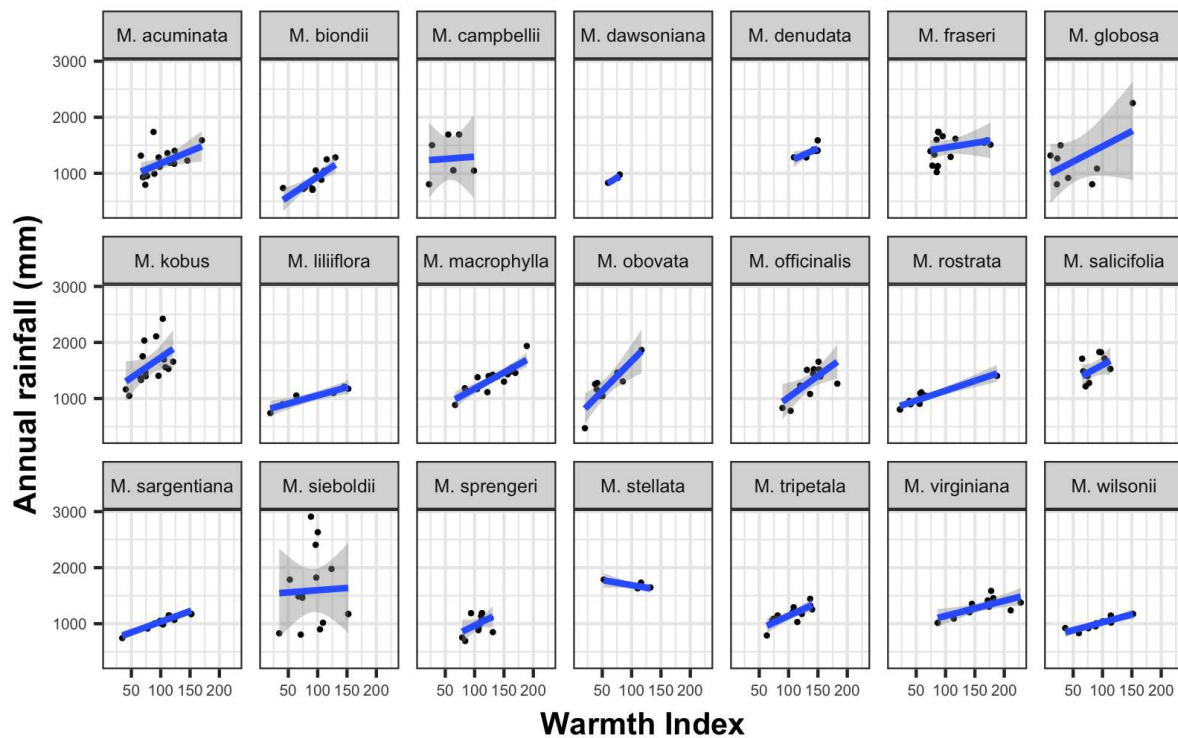
326  $r = 0.3161698$ ,  $p \leq 0.000001$ ,  $n = 248$ ,  $R^2 = 0.09629$

327

328

329 On the basis of the weak inter-specific variation found in the heuristic literature, further analysis was  
 330 only carried out for the observational literature. To explore these inter-specific differences, SDMs for  
 331 each species were created, again plotting Annual rainfall against Warmth Index in Figure 5. In these  
 332 models, intercept and slope vary, as does the degree of fit between the regression line and the  
 333 distribution of populations, with species such as *M. globosa* and *M. sieboldii* showing eurytopic  
 334 tolerances (i.e., an ability to adapt to a wide range of conditions) and *M. liliiflora*, *M. rostrata* and *M.*  
 335 *sargentiana* showing stenotopic behaviours (i.e., occupation of a restricted range of conditions).

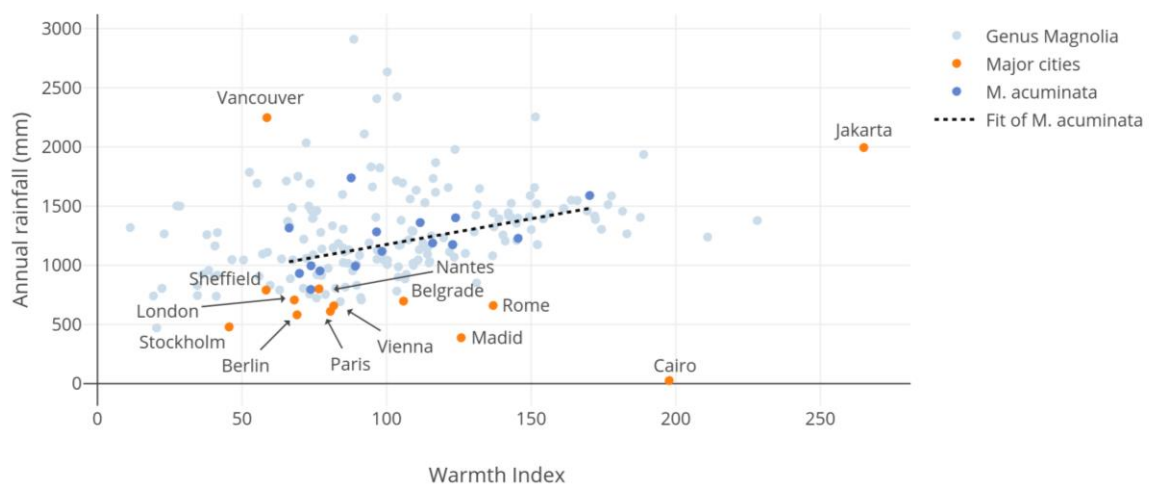
336  
 337 *Figures 5. Intraspecific variation in climate niche in 21 wild-growing, deciduous Magnolia species, as reported by the*  
 338 *observational literature*



339  
 340 *Identifying Magnolia populations suitable for urban forestry under climate change*  
 341 In the final step of the analysis, selected species that showed a range of gradients, intercepts and fit  
 342 were re-plotted in the context of cities that represent a range of urban forestry conditions (Figs 6a-c).  
 343 In these figures the regression line of the *Magnolia* species distribution allows hypothetical  
 344 fundamental niches to be identified and compared with conditions currently experienced by major  
 345 cities. This study found that the regression line in certain species is very close to environments found  
 346 in European cities (e.g. *M. biondii*, *M. officinalis*), suggesting that they would be better fitted to urban

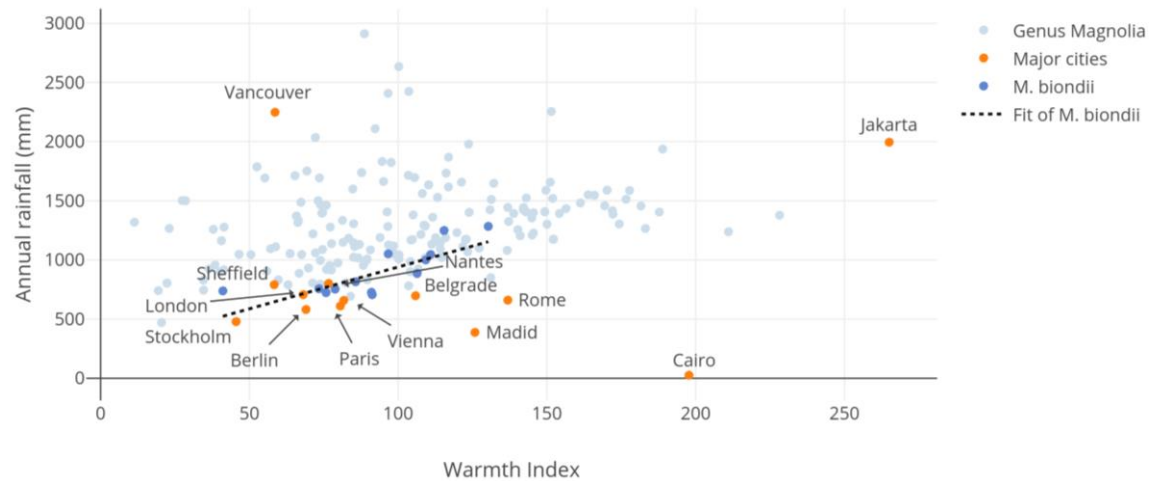
347 forestry applications in some cities than others, and that within these species, certain populations are  
 348 likely to be particularly well fitted. Nevertheless, even within species that do not show a regression  
 349 line closely intersecting with some cities, it is not uncommon to find outlier populations that may be  
 350 appropriate. Across most species it was found that while there was often an overlap between the  
 351 Warmth Index in northern European cities and naturally distributed *Magnolia* populations, there was  
 352 typically a shortfall of rainfall in the urban situations, which would create an imbalance in the water-  
 353 energy relations in urban forestry.

354 *Figure 6a. The climate niches experienced by M. acuminata in relation to those found in selected major cities*



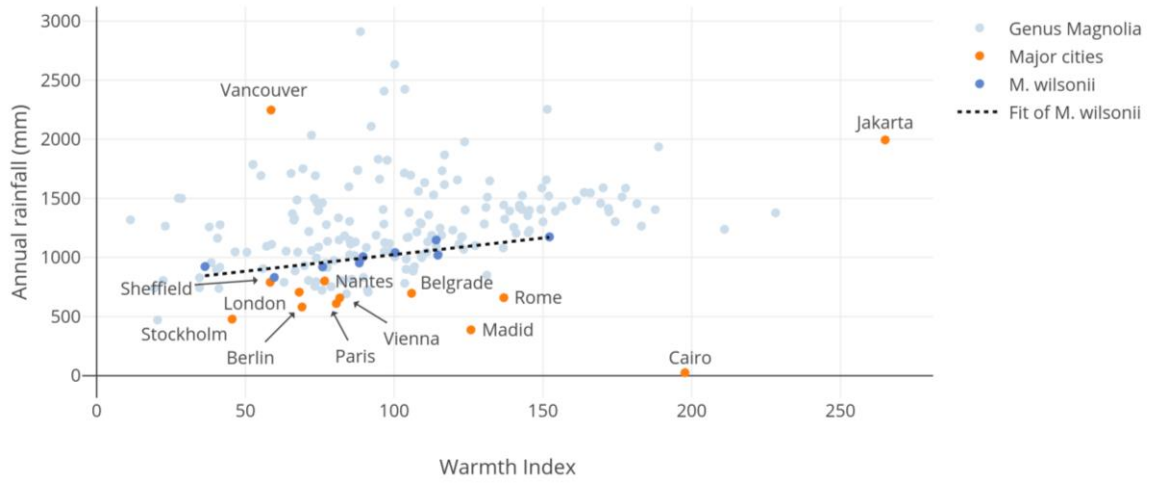
355

356 *Figure 6b. The climate niches experienced by M. biondii in relation to those found in selected major cities*



357

358 *Figure 6c. The climate niches experienced by M. wilsonii in relation to those found in selected major cities*



359

## 360 **Discussion**

### 361 *Clear differences between literature sources*

362 The results of this literature review demonstrate clear differences between the three literature sources  
363 (see Table 6). The experimental literature is highly focussed, offering insights into specific  
364 physiological traits and at times, intra-specific variation in response to stresses, but whilst it may yet  
365 be possible to create a comprehensive understanding of how functional ecology concepts might  
366 influence urban forestry in some genera, this is not currently possible for the genus *Magnolia* given  
367 the small number of studies that have been carried out. In addition, there remain numerous  
368 methodological issues with recording functional traits that may complicate specification for urban  
369 foresters, primarily that individual traits are not significant unless they are contextualised either  
370 against other species, or are shown to trade-off against other traits within the same species. Further  
371 inherent complications exist with using traits presented in large datasets, such as TRY (Kattge *et al.*,  
372 2011) or Bien R (Maitner *et al.*, 2018), given that the reported traits have been collected on different  
373 individuals under different conditions (e.g. some under manipulated conditions (Toledo-Aceves,  
374 López-Barrera and Vásquez-Reyes, 2017), others in common garden experiments whilst yet others are  
375 gathered in the wild). Other studies in this literature present findings from trait studies but not the data  
376 themselves, making them impossible to interrogate or contextualise (Kitaoka *et al.*, 2016; Oguchi,  
377 Hiura and Hikosaka, 2017). Attempts to explain trait coordination through strategies or Plant  
378 Functional Types are highly attractive and hold great promise for specification in urban forestry but at  
379 the time of writing, remain elusive: conceptually elegant schemes such as the Fast Slow Spectrum  
380 (Reich, 2014) or the CSR triangle (Pierce *et al.*, 2017) rely upon proxy traits to too great an extent at  
381 present to explain functional trait coordination within woody plants. It is not yet clear how leaf data,  
382 for example, can be reliable proxies for reproductive traits (which are highly important sources of  
383 photosynthetic investment in *Magnolia*), nor how leaf turgor loss point (as reported by Sjöman *et al.*  
384 (2018)) is traded-off against other traits. Whilst a great deal of research has been carried out in these  
385 areas for forestry trees, species appropriate for urban forestry and horticulture have not been assessed  
386 using the same methodologies and bridging this gap should be a priority for researchers. Perhaps most

387 problematic for this literature is the fact that although relationships between tolerance of stresses such  
388 as drought, soil salinity or winter cold have been hypothesised in woody plants and experimentally  
389 shown in herbaceous plants (Grime et al. 2007), this relationship has yet to be resolved and does not  
390 in itself help urban foresters infer resource demand or optimal opportunities for species selection in  
391 urban forest.

392

393 The heuristic literature is extensive, but in most cases rather superficial, typically giving an overview  
394 of the genus *Magnolia*, making it difficult to compare the qualities of different species let alone the  
395 different characteristics of populations found within a species. Much of the focus in this literature is  
396 on ornamental qualities rather than functional or ecological aspects, making it difficult to accurately  
397 assess the likely fit between a species and a planting situation in urban forestry. Further, the lack of  
398 consistent vocabulary across these sources (Table 4) means that factors such as reported hardiness are  
399 difficult to interpret consistently, not only because of the variation in hardiness schemes used (this  
400 might be either the USDA or RHS hardiness ratings, although some sources discuss hardiness in  
401 terms of tolerance of other stresses such as soil alkalinity) but also because different standards have  
402 been applied to categorise plants within the scheme and often without using standardised trials.

403

404 The observational literature illustrates many of the challenges of using big data to answer practical  
405 challenges, the first being the quality of the data that is used. Gathering the records for the plant  
406 occurrence data was a long-winded process, with each of the large databases presenting their own  
407 challenges: GBIF, for example, holds relatively few verified observations of naturally occurring  
408 *Magnolia* populations but offers excellent data transfer capabilities, Global Plants hosts a large  
409 number of records but makes data transfer challenging, whilst the Chinese Virtual Herbarium requires  
410 translation from Mandarin and an iterative process of positive identification and filtering to derive  
411 accurate records. The archival research in herbaria was highly effective but corroborated reports that  
412 only a small fraction of plant records are hosted by online databases (Harris and Marsico, 2017;  
413 Kirchhoff *et al.*, 2018), and as a result future applications of this methodology should factor in the  
414 extensive desktop research. By contrast, climate data was straightforward to derive, with the principal

415 short-comings being the grain and accuracy of the data, especially when assessing urban  
416 environments in comparison to their rural hinterlands. Given that urban forestry sites are typically  
417 affected by urban heat island effects in combination with localised variation in solar reflectivity, salt  
418 spray and soil compaction, this methodology should perhaps be seen as a framework which can guide  
419 experimental research rather than a stand-alone decision making tool. Interpolation of climate data for  
420 locations between climate stations allows for an estimate to be generated for any given location but it  
421 is not clear to what extent these models account for elevation, aspect, slope or surface features, i.e.  
422 factors that affect temperature and vapour pressure deficit, which is particularly important in  
423 mountainous areas. Nevertheless, temperature and precipitation have been demonstrated to be key  
424 determinants of a Grinnellian niche (i.e., the effect of the environment on species distribution (Gravel  
425 *et al.*, 2018)) and should data become more nuanced, readily available or easier to use, this  
426 methodology would have strong potential for urban forestry specification as it appears to reveal not  
427 only intra-specific variation but also the climate niche occupied by various populations, thereby  
428 making it possible to match naturally-distributed populations to actual planting locations in designed  
429 landscapes. It is not clear at this stage whether the variation found at intraspecific levels are a result of  
430 genetic variation, other factors such as Cold Index, timing of resource availability, edaphic factors or  
431 cultural processes, or simply due to species occupying ranges outside their fundamental niche but this  
432 might also be resolved through more complex studies using hierarchical framework models or more  
433 data. From an urban forestry perspective, as opposed to a horticultural process where greater  
434 management resources are available, the tipping points that trigger mortality or poor performance are  
435 essential to understand and whilst this literature has the potential to explain environmental resource  
436 availability, it does not yet reveal the thresholds for fatal decline that are triggered by phenomena  
437 such as aseasonal drought or extended periods of anoxia due to soil flooding or mechanical  
438 compaction.

439 *Table 6. Efficacy of literature sources for specifying diverse ecotypes of trees in urban forestry*

	<b>Heuristic literature</b>	<b>Experimental literature</b>	<b>Observational literature</b>
Discusses broad range of species	Yes	Yes	Yes

Identifies intra-specific genetic variation	No	Potentially	Yes
Effective at identifying urban-fit species	No	Potentially	Yes
Effective at screening for new ecotypes	No	No	Yes

440

441 *Species Distribution Models and ecotype matching*

442 The big data held in herbaria and in climate models offer tantalising opportunities to improve urban  
443 forestry specification but assessing these data in SDMs needs careful examination to understand their  
444 implications. Fig 7 uses the data from *M. obovata* and *M. officinalis*, two closely related species that  
445 are often seen as having similar horticultural requirements, to identify four key concepts that  
446 demonstrate the practical applications of using SDMs in ecotype selection: the gradient of the slope  
447 (A) describes the underlying metabolism of the species, showing how constrained water – energy  
448 relations are: a shallow gradient, for example, would indicate that warmth index (energy) is not a  
449 constraint upon photosynthesis whilst a steep gradient would indicate that it is a critical factor for  
450 growth. Comparing the lengths of the regression lines (B) in each species allows us to ask ‘what are  
451 the factors that determine the start and end points of slope, and thus limit the distribution of  
452 populations within the species?’ These limiting factors are likely to be different at each end of the  
453 regression line – evapotranspiration may be too great in ranges with high annual rainfall and warmth  
454 index, for example, whilst insufficient solar radiation or temperature during the growing season may  
455 prevent some populations from creating enough lignin during the growing season to tolerate winter  
456 cold. Further, these limiting factors do not need to be lethal in order to be effective, rather they may  
457 be just enough to stop physiological or reproductive processes from being sufficiently effective to  
458 ensure species range extension. The location of the intercept (C) on the other hand, indicates the  
459 relative effect of water as a constraint upon growth, with intercepts higher up the y axis indicating  
460 increasing importance of this resource. The 95% confidence interval (D) the regression line indicates  
461 the degree of variability between the samples and can be used to assess the robustness of the data. The

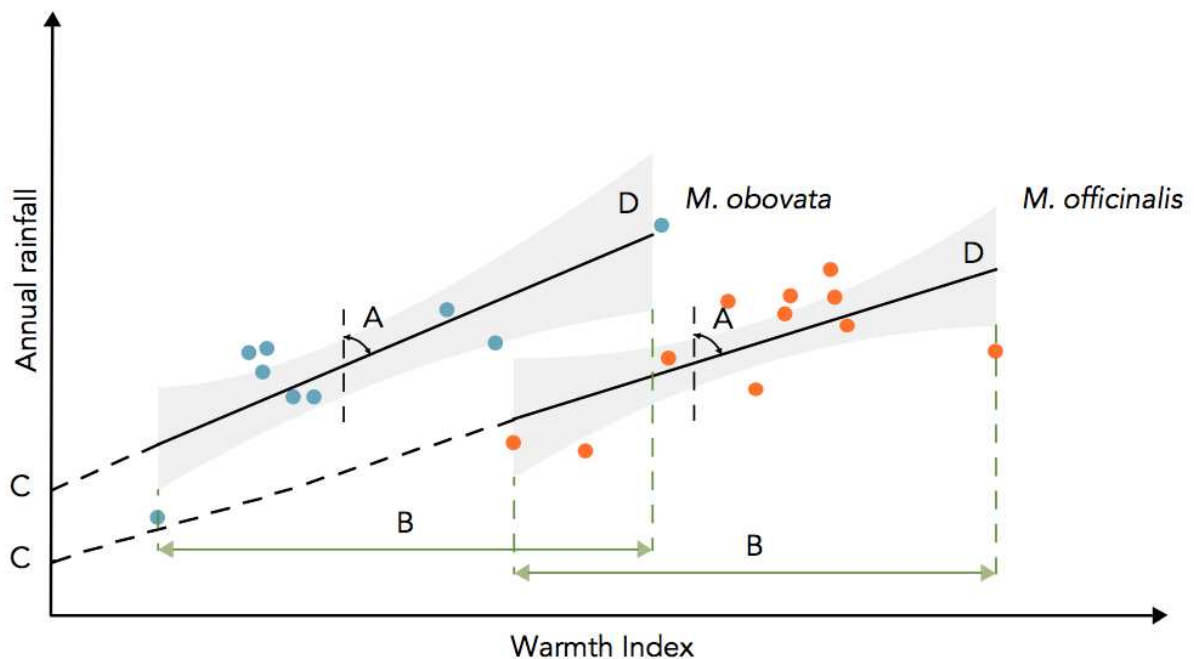
462 sum of these subtle differences demonstrates that the two species occupy two different niches, with  
463 *M. obovata* distributed in ranges with more rainfall per unit of warmth index than *M. officinalis*,  
464 suggesting that *M. obovata* may be more water-demanding than *M. officinalis*.

465

466 Future studies should test these hypotheses using other bioclimatic variables such as soil pH, soil  
467 oxygen or community-level factors, with trials to test the thresholds for mortality under stress. Such  
468 studies would also be able to answer questions of whether the degree of variance from the regression  
469 line corresponds to geographic range and whether the regression line corresponds to abundance  
470 models along gradients of physical environmental conditions (Cox and Moore 2010).

471 *Figure 7. A conceptual implications of Species Distribution Models for urban foresters*

472



473

474

475 ***Intra-specific variation is revealed in Species Distribution Models.***

476 Cox and Moore (2010) argue that given the climatic fluctuations of past 2 million years, extant  
477 *Magnolia* species are likely to be the most competitive species in the genus's history: species with  
478 older phylogenies were often too small and slow-growing to compete with faster growing species as  
479 the planet warmed, shaping the possibilities for future evolutionary outcomes. As a result, the range of  
480 traits possessed by *Magnolia* species that we observe today are unusually conserved and may, in  
481 relation to other genera, present a picture of relatively narrow variation. Nevertheless, as reported in

482 *Quercus* (Barbero, Loisel and Quézel, 1992; García-Nogales *et al.*, 2016) and *Nothofagus* (Fajardo  
483 and Piper, 2011; Richardson *et al.*, 2013), trees have remarkable capacity to adjust their physiology  
484 and morphology to different climates and that these adjustments can be explained either through  
485 varying evolutionary strategies to tolerate stress (Grime and Pierce, 2012) or sub-specific / population  
486 level genotype or phenotype variation. Understanding the source and level of variation is likely to be  
487 critical to successful urban forestry specification and it appears that when combined with target design  
488 sites, these basic SDMs are capable of identifying populations of particular relevance both to  
489 specifiers who wish to select plants from a particular provenance or for producers who wish to  
490 identify populations with particular promise for breeding or selection studies. The ‘stable rear edge’  
491 of a population identified by (Hampe and Petit, 2005) can be located using these SDMs, making it  
492 possible to identify populations with higher levels of genetic diversity- and conversely, leading edges  
493 of a population with reduced diversity and therefore a greater probability of possessing specific traits.

494

#### 495 *Water - energy relations appear to be important drivers of species distribution- and trait variation*

496 Although current climate is not the only factor that affect species distribution or genetic diversity, the  
497 energy hypothesis proposed by Hawkins *et al.* (2003) provides a compelling explanation for the  
498 distribution of species within two axes of variation. Water-use strategies have been shown to be  
499 related to environmental conditions (Baastrup-Spohr *et al.*, 2015; Aguilar-Romero *et al.*, 2017) and  
500 the distribution models in Figures 7 and 8 illustrate these mechanisms in genus *Magnolia*, supporting  
501 the proposal by Hawkins *et al.* (2003) that water variables tend to be stronger predictors in sub-  
502 tropical and warm temperate climates, whilst water-energy variables tend to be stronger predictors in  
503 cold temperate regions. There appears to be a strong consensus that climate is a significant  
504 determinant of a species range (Normand *et al.*, 2011), with plant trait variation associated with  
505 adaption to light and water availability, and a coordinated tolerance of plants to shortages of both  
506 resources proposed (Cavender-Bares, Kitajima and Bazzaz, 2004; Castellanos-Castro and Newton,  
507 2015).

508

509 Based upon this understanding of water-energy relations, ecologists are modelling environmental  
510 niches from traits (Cadotte *et al.*, 2015) but the question for urban foresters is whether the reverse can  
511 be modelled, i.e., whether we can speculate that traits vary in accordance to climate niche: this  
512 process is well established in commercial forestry in Western Europe, drawing on decades of spruce  
513 evaluation but this process is not as developed in broadleaf woody plants. Cavender-Bares *et al.*  
514 (2004) argued that phenotype specialisation explains niche adaptation but it is not yet clear to what  
515 extent phenotype traits are inheritable or under epigenetic control. Carmona *et al.* (2016) offer a  
516 potential methodology for resolving using highly complex models that require higher levels of  
517 sophistication and data than assessed in this study; similarly, advances in molecular ecology establish  
518 links between populations and traits, with Beaulieu *et al.* (2011) finding relationships between  
519 phenotype SNPs and traits, paving the way for marker-assisted selection in tree species.

520

#### 521 *What have we learnt about using Magnolia species in urban forestry?*

522 Urban foresters wish to maximise the fit between trees and their environment and typically this means  
523 knowing whether some species are better suited to certain roles than others: by using the results set  
524 out in Fig 6 it is possible to hypothesise that *M. biondii* and *M. wilsonii*, for example, are likely to be  
525 well-suited to use in north-western European urban forestry due to the current close overlap between  
526 the climatic conditions in their natural habitat and cities in these locations, although this might be  
527 expected to evolve under climate change. Designers might use this information to select *M. biondii* as  
528 street trees and *M. wilsonii* in situations where shrubbier forms are more appropriate such as stylised-  
529 coppice communities: by contrast, it appears that whilst *M. acuminata* displays cold-tolerance, water  
530 availability is likely to be an important factor in determining fit and as such, these species might be  
531 more appropriate in SuDS environments where a greater water availability can be designed. Most  
532 importantly, this research shows it is possible to specify *Magnolia* in urban forestry with much greater  
533 precision than the ‘species’ level, allowing us to identify alternative species or ecotypes based on the  
534 constraints of a given location, accounting for micro-climatic variations due to factors such as the  
535 albedo effect (which increases evapotranspiration) or SuDS design (which would increase available

536 water in the root zone). The findings demonstrate not only that there are a range of niches occupied by  
537 each species (and therefore some degree of niche adaptation) but that the tools to identify these niches  
538 and match them to existing and future urban environments exist. This finding creates exciting  
539 opportunities for collection strategies and the introduction of new genetic material to horticulture  
540 (Kardos and Shafer, 2018).

541

#### 542 *Future applications and further studies*

543 Following early attempts to use biogeography concepts to specify street trees (Jim, 1988), the  
544 availability of large data sets of plant occurrences and climate open new opportunities for urban  
545 foresters to reinvigorate this area of research, building upon well-established biogeographical  
546 practices (Elith and Leathwick, 2009). Indeed, alongside improving specification practices, urban  
547 foresters could become part of climate adaptation and assisted migration strategies if provenance  
548 identification and ecotype selection were developed (Fontaine and Larson, 2016). Given the specialist  
549 skills required to gather and interpret the necessary data, in the first instance it should be possible to  
550 create a proof-of-concept website that pulls plant occurrence data from online data repositories (such  
551 as GBIF) and uses a pivot table to interact with climate data, and then perform the basic mathematics  
552 to produce a basic SDM for a given species in relation to a urban forestry target sites: although these  
553 graphs would draw upon a limited number of occurrences, such an application would powerfully  
554 illustrate the capabilities of this line of research and rapidly identify knowledge gaps in other genera.

555

556 A similarly important step would be to assess whether climatic factors have the same degree of  
557 explanatory power in both natural and designed environments: whilst water-energy balances might  
558 account for the greatest degree of fit in 95% cases of naturally-distributed plants (Hawkins *et al.*,  
559 2003), factors such as soil anaerobia, compaction, pollution or disturbance would be expected to play  
560 significant roles in urban environments. These relationships could be tested through further desktop  
561 studies, using complex hull analysis to incorporate soils data and traits (where available) or in  
562 common garden experiments, examining chloroplast and carbon allocation through time under a range  
563 of stressful conditions would allow the limits of big data's utility to urban foresters to be explored.



## 565 **Conclusion**

566 In this paper we review the literature that is available to urban foresters to specify trees and  
567 quantitatively review the aims and reliability of the sources. Using *Magnolia* as a case-study genus,  
568 we find that there are considerable differences between the literatures, ranging from broad stereotypes  
569 of ideal growing environments in the heuristic literature to highly precise, non-contextualised studies  
570 examining single traits within a species in the experimental literature. Whilst the experimental  
571 literature provides a gold-standard of evidence for understanding plant functioning, only a small  
572 proportion of the plants that urban foresters might be interested in have been studied and most of  
573 these studies are reported in academic journals, resending barriers to access. In spite of its short-  
574 comings the heuristic literature is therefore the first port of call for most practitioners, with the result  
575 that they are unlikely to specify novel species or provenances with confidence.

576

577 To overcome this, we identify a new literature source and develop a methodology for ecotype  
578 selection that could be used both by urban forestry researchers and the nursery trade, drawing upon  
579 well-established biogeographical theory and big data. The development and availability of big data  
580 allows urban foresters to harness biogeographical techniques, combining precise, quantitative  
581 empirical studies within a holistic understanding of plant-environment relations. Whilst this approach  
582 requires further testing in other genera and testing against other variables that affect species  
583 distribution and fit, using species distribution modelling holds considerable promise for recognising  
584 the fundamental distinction between preferred growing conditions and the environmental limits that  
585 trees can withstand, and developing urban forestry discourse and practice.

## 586 **Acknowledgements**

587 This research did not receive any specific grant from funding agencies in the public, commercial, or  
588 not-for-profit sectors.

589

## 590 **Bibliography**

- 591 Acuto, M. *et al.* (2018) *Science and the Future of Cities*. London and Melbourne. doi:  
592 10.13140/RG.2.2.27706.64969.
- 593 Aguilar-Romero, R. *et al.* (2017) ‘Differentiation in the water-use strategies among oak species from  
594 central Mexico’, *Tree physiology*, 37, pp. 915–925. doi: 10.1093/treephys/tpx033.
- 595 Allen, C. D., Breshears, D. D. and McDowell, N. G. (2015) ‘On underestimation of global  
596 vulnerability to tree mortality and forest die-off from hotter drought in the Anthropocene’, *Ecosphere*,  
597 6(8), pp. 1–55. doi: 10.1890/ES15-00203.1.
- 598 Allen, J. M. *et al.* (2018) ‘Biodiversity synthesis across the green branches of the tree of life’, *Nature*  
599 *Plants*. Springer US. doi: 10.1038/s41477-018-0322-7.
- 600 Azuma, H. *et al.* (2001) ‘Molecular phylogeny of the Magnoliaceae: The biogeography of tropical  
601 and temperate disjunctions’, *American Journal of Botany*, 88(8812), pp. 2275–2285. Available at:  
602 <http://www.jstor.org/stable/3558389>.
- 603 Azuma, H. *et al.* (2011) ‘Intraspecific Sequence Variation of cpDNA Shows Two Distinct Groups  
604 Within *Magnolia virginiana* L. of Eastern North America and Cuba’, *Castanea*, 76(1), pp. 118–123.  
605 Available at: <http://www.jstor.org/stable/41301586>.
- 606 Azuma, H., Toyota, M. and Asakawa, Y. (2001) ‘Intraspecific Variation of Floral Scent Chemistry in  
607 *Magnolia kobus* DC. (Magnoliaceae)’, *Journal of Plant Research*. doi: 10.1007/PL00014006.
- 608 Baastrup-Spohr, L. *et al.* (2015) ‘From soaking wet to bone dry: Predicting plant community  
609 composition along a steep hydrological gradient’, *Journal of Vegetation Science*, 26, pp. 619–630.  
610 doi: 10.1111/jvs.12280.
- 611 Barbero, M., Loisel, R. and Quézel, P. (1992) ‘Biogeography, ecology and history of Mediterranean  
612 *Quercus ilex* ecosystems’, *Vegetatio*, 99, pp. 19–34. doi: 10.1007/BF00118207.
- 613 Barbrook, J. *et al.* (2018) *The Right Tree in the Right Place for a Resilient Future*.
- 614 Bassuk, N. L. *et al.* (2009) *Recommended Urban Trees: Site Assessment and Tree Selection for Stress*  
615 *Tolerance*. Urban Horticulture Institute.
- 616 Beaulieu, J. *et al.* (2011) ‘Association genetics of wood physical traits in the conifer white spruce and

617 relationships with gene expression’, *Genetics*, 188, pp. 197–214. doi: 10.1534/genetics.110.125781.

618 Bialecki, M. B., Fahey, R. T. and Scharenbroch, B. (2018) ‘Variation in urban forest productivity and  
619 response to extreme drought across a large metropolitan region’, *Urban Ecosystems*. *Urban*  
620 *Ecosystems*, 21(1), pp. 157–169. doi: 10.1007/s11252-017-0692-z.

621 Booth, T. H. (2018) ‘Species distribution modelling tools and databases to assist managing forests  
622 under climate change’, *Forest Ecology and Management*, 430, pp. 196–203. doi:  
623 10.1016/j.foreco.2018.08.019.

624 Budd, C., Zimmer, E. and Freeland, J. R. (2015) ‘Conservation genetics of *Magnolia acuminata*, an  
625 endangered species in Canada: Can genetic diversity be maintained in fragmented, peripheral  
626 populations?’, *Conservation Genetics*, 16, pp. 1359–1373. doi: 10.1007/s10592-015-0746-9.

627 Cadotte, M. W. *et al.* (2015) ‘Predicting communities from functional traits’, *Trends in Ecology and*  
628 *Evolution*, 30(9), pp. 510–511. doi: 10.1016/j.tree.2015.07.001.

629 Carmona, C. P. *et al.* (2016) ‘Traits Without Borders: Integrating Functional Diversity Across Scales’,  
630 *Trends in Ecology and Evolution*, 31(5), pp. 382–394. doi: 10.1016/j.tree.2016.02.003.

631 Castellanos-Castro, C. and Newton, A. C. (2015) ‘Leaf and stem trait variation and plant functional  
632 types in 113 woody species of a seasonally dry tropical forest’, *Colombia Forestal*, 18(50), pp. 117–  
633 138. doi: 10.14483/udistrital.jour.colomb.for.2015.1.a07.

634 Cavender-Bares, J., Kitajima, K. and Bazzaz, F. A. (2004) ‘Multiple Trait Associations in Relation to  
635 Habitat Differentiation among 17 Floridian Oak Species’, *Ecological Monographs*, 74(4), pp. 635–  
636 662. doi: 10.1890/03-4007.

637 Cicuzza, D., Newton, A. and Oldfield, S. (2007) *The Red List of Magnoliaceae*. Cambridge: Flora &  
638 Fauna International.

639 Cires, E. *et al.* (2013) ‘Gap analyses to support ex situ conservation of genetic diversity in *Magnolia*,  
640 a flagship group’, *Biodiversity and Conservation*, 22, pp. 567–590. doi: 10.1007/s10531-013-0450-3.

641 Díaz, S. *et al.* (2016) ‘The global spectrum of plant form and function’, *Nature*, 529, pp. 167–171.  
642 doi: 10.1038/nature16489.

643 Elith, J. and Leathwick, J. R. (2009) ‘Species Distribution Models: Ecological Explanation and  
644 Prediction Across Space and Time’, *Annual Review of Ecology, Evolution, and Systematics*, 40(1), pp.

645 677–697. doi: 10.1146/annurev.ecolsys.110308.120159.

646 Fajardo, A. and Piper, F. I. (2011) ‘Intraspecific trait variation and covariation in a widespread tree  
647 species (*Nothofagus pumilio*) in southern Chile’, *New Phytologist*, 189, pp. 259–271. doi:  
648 10.1111/j.1469-8137.2010.03468.x.

649 Fontaine, L. C. and Larson, B. M. H. (2016) ‘The right tree at the right place? Exploring urban  
650 foresters’ perceptions of assisted migration’, *Urban Forestry and Urban Greening*, 18, pp. 221–227.  
651 doi: 10.1016/j.ufug.2016.06.010.

652 Fuller, L. and Quine, C. P. (2016) ‘Resilience and tree health: A basis for implementation in  
653 sustainable forest management’, *Forestry*, 89, pp. 7–19. doi: 10.1093/forestry/cpv046.

654 García-Nogales, A. *et al.* (2016) ‘Range-wide variation in life-history phenotypes: spatiotemporal  
655 plasticity across the latitudinal gradient of the evergreen oak *Quercus ilex*’, *Journal of Biogeography*,  
656 43, pp. 2366–2379. doi: 10.1111/jbi.12849.

657 Gravel, D. *et al.* (2018) ‘Bringing Elton and Grinnell together: a quantitative framework to represent  
658 the biogeography of ecological interaction networks’, *Ecography*, 41, pp. 1–15. doi:  
659 10.1111/ecog.04006.

660 Hallgren, W. *et al.* (2016) ‘The Biodiversity and Climate Change Virtual Laboratory: Where ecology  
661 meets big data’, *Environmental Modelling and Software*. Elsevier Ltd, 76, pp. 182–186. doi:  
662 10.1016/j.envsoft.2015.10.025.

663 Hampe, A. and Petit, R. J. (2005) ‘Conserving biodiversity under climate change: the rear edge  
664 matters’, *Ecology Letters*, 8(5), pp. 461–467.

665 Harris, K. M. and Marsico, T. D. (2017) ‘Digitizing Specimens in a Small Herbarium: A Viable  
666 Workflow for Collections Working with Limited Resources’, *Applications in Plant Sciences*, 5(4), p.  
667 1600125. doi: 10.3732/apps.1600125.

668 Hawkins, B. A. *et al.* (2003) ‘Energy, water, and broad-scale geographic patterns of species richness’,  
669 (July 2003), pp. 3105–3117. doi: 10.1890/03-8024.

670 Haxeltine, A. and Prentice, I. C. (1996) ‘An equilibrium terrestrial biosphere model based on  
671 ecophysiological constraints, resource availability, and competition among plant functional types’,  
672 *Global biogeochemical cycles*, 10(4), pp. 693–709.

673 Hirons, A. D. and Sjoman, H. (2018) *Tree species selection for Green Infrastructure: a guide for*  
674 *specifiers*.

675 Hitchmough, J. D. and Bonugli, A. M. (1997) 'Attitudes of residents of a medium sized town in south  
676 west Scotland to street trees', *Landscape Research*, 22(3), pp. 327–337. doi:  
677 10.1080/01426399708706518.

678 Hoyle, H., Hitchmough, J. and Jorgensen, A. (2017) 'Attractive, climate-adapted and sustainable?  
679 Public perception of non-native planting in the designed urban landscape', *Landscape and Urban*  
680 *Planning*. Elsevier B.V., 164, pp. 49–63. doi: 10.1016/j.landurbplan.2017.03.009.

681 Ippoliti, E. (2015) *Heuristic Reasoning: Studies in Applied Philosophy, Epistemology and Rational*  
682 *Ethics*. Switzerland: Springer International Publishing.

683 James, S. A. *et al.* (2018) 'Herbarium data: Global biodiversity and societal botanical needs for novel  
684 research', *Applications in Plant Sciences*, 6(2), pp. 1–8. doi: 10.1002/aps3.1024.

685 Jim, C. Y. (1988) 'Street tree study as a theme in urban biogeography', *Geography*, 73(3), pp. 226–  
686 232.

687 Kardos, M. and Shafer, A. B. A. (2018) 'The Peril of Gene-Targeted Conservation', *Trends in*  
688 *Ecology and Evolution*. doi: 10.1016/j.tree.2018.08.011.

689 Kattge, J. *et al.* (2011) 'TRY - a global database of plant traits', *Global Change Biology*, 17(9), pp.  
690 2905–2935. doi: 10.1111/j.1365-2486.2011.02451.x.

691 Kirchhoff, A. *et al.* (2018) 'Toward a service-based workflow for automated information extraction  
692 from herbarium specimens', *Database : the journal of biological databases and curation*, 2018, pp.  
693 1–11. doi: 10.1093/database/bay103.

694 Kitaoka, S. *et al.* (2016) 'The photosynthetic response of four seral deciduous broad-leaved tree  
695 seedlings grown under elevated CO2 concentrations', *Journal of Agricultural Meteorology*, 72(1), pp.  
696 43–49. doi: 10.2480/agrmet.D-14-00016.

697 Krajter Ostoić, S. and Konijnendijk van den Bosch, C. C. (2015) 'Exploring global scientific  
698 discourses on urban forestry', *Urban Forestry and Urban Greening*, 14, pp. 129–138. doi:  
699 10.1016/j.ufug.2015.01.001.

700 Kreft, H. and Jetz, W. (2007) 'Global patterns and determinants of vascular plant diversity', 104(14),

701 pp. 5925–5930. doi: 10.1073/pnas.0608361104.

702 Maitner, B. S. *et al.* (2018) ‘The bien r package: A tool to access the Botanical Information and  
703 Ecology Network (BIEN) database’, *Methods in Ecology and Evolution*, 9(2), pp. 373–379. doi:  
704 10.1111/2041-210X.12861.

705 Morgenroth, J. *et al.* (2016) ‘Urban tree diversity-Taking stock and looking ahead’, *Urban Forestry  
706 and Urban Greening*, 15, pp. 1–5. doi: 10.1016/j.ufug.2015.11.003.

707 Muranishi, S. *et al.* (2013) ‘Asymmetric introgression between *Magnolia stellata* and *M. salicifolia* at  
708 a site where the two species grow sympatrically’, *Tree Genetics and Genomes*, 9, pp. 1005–1015. doi:  
709 10.1007/s11295-013-0612-1.

710 Normand, S. *et al.* (2011) ‘Postglacial migration supplements climate in determining plant species  
711 ranges in Europe’, *Proceedings of the Royal Society B: Biological Sciences*, 278(1725), pp. 3644–  
712 3653. doi: 10.1098/rspb.2010.2769.

713 Oguchi, R., Hiura, T. and Hikosaka, K. (2017) ‘The effect of interspecific variation in photosynthetic  
714 plasticity on 4-year growth rate and 8-year survival of understory tree seedlings in response to gap  
715 formations in a cool-temperate deciduous forest’, *Tree Physiology*, 37(8), pp. 1113–1127. doi:  
716 10.1093/treephys/tpx042.

717 Ordóñez Barona, C. (2015) ‘Adopting public values and climate change adaptation strategies in urban  
718 forest management: A review and analysis of the relevant literature’, *Journal of Environmental  
719 Management*, 164, pp. 215–221. doi: 10.1016/j.jenvman.2015.09.004.

720 Pelletier, T. A. *et al.* (2018) ‘Predicting plant conservation priorities on a global scale’, *Proceedings  
721 of the National Academy of Sciences*, 115(51), p. 201804098. doi: 10.1073/pnas.1804098115.

722 Pierce, S. *et al.* (2017) ‘A global method for calculating plant CSR ecological strategies applied  
723 across biomes world-wide’, *Functional Ecology*, 31, pp. 444–457. doi: 10.1111/1365-2435.12722.

724 Reich, P. B. (2014) ‘The world-wide “fast-slow” plant economics spectrum: A traits manifesto’,  
725 *Journal of Ecology*. doi: 10.1111/1365-2745.12211.

726 Richardson, S. J. *et al.* (2013) ‘Intraspecific Relationships among Wood Density, Leaf Structural  
727 Traits and Environment in Four Co-Occurring Species of *Nothofagus* in New Zealand’, *PLoS ONE*,  
728 8(3), pp. 24–26. doi: 10.1371/journal.pone.0058878.

729 Rivers, M. *et al.* (2016) *The Red List of Magnoliaceae*. London: Botanic Gardens Conservation  
730 International.

731 Roloff, A., Korn, S. and Å, S. G. (2009) ‘The Climate-Species-Matrix to select tree species for urban  
732 habitats considering climate change’, *Urban Forestry & Urban Greening*. Elsevier, 8(4), pp. 295–  
733 308. doi: 10.1016/j.ufug.2009.08.002.

734 Samson, R. *et al.* (2017) ‘Species-Specific Information for Enhancing Ecosystem Services’, in  
735 Pearlmutter, D. *et al.* (eds) *The Urban Forest*. Heidelberg: Springer Nature, pp. 114–144.

736 Santamour, F. S. (1990) ‘Trees for urban planting : diversity, uniformity , and common sense’,  
737 *Agriculture*, 7, pp. 57–65. Available at:  
738 [http://books.google.com/books?hl=en&lr=&id=SAJQhK1fWDoC&oi=fnd&pg=PA396&dq=Trees+for+urban+planting:+Diversity,+uniformity,+and+common+sense&ots=RklXdDqjbJ&sig=7QcH6Fg\\_aOv-bY6DefzVboGZBAM](http://books.google.com/books?hl=en&lr=&id=SAJQhK1fWDoC&oi=fnd&pg=PA396&dq=Trees+for+urban+planting:+Diversity,+uniformity,+and+common+sense&ots=RklXdDqjbJ&sig=7QcH6Fg_aOv-bY6DefzVboGZBAM).  
740

741 Serra-Diaz, J. M. *et al.* (2017) ‘Big data of tree species distributions: how big and how good?’, *Forest  
742 Ecosystems*. Forest Ecosystems, 4(1). doi: 10.1186/s40663-017-0120-0.

743 Sjöman, H., Hirons, A. D. and Bassuk, N. L. (2018) ‘Magnolias as urban trees—a preliminary  
744 evaluation of drought tolerance in seven magnolia species’, *Arboricultural Journal*. Taylor & Francis,  
745 40(1), pp. 47–56. doi: 10.1080/03071375.2017.1415554.

746 Toledo-Aceves, T., López-Barrera, F. and Vásquez-Reyes, V. (2017) ‘Preliminary analysis of  
747 functional traits in cloud forest tree seedlings’, *Trees - Structure and Function*. Springer Berlin  
748 Heidelberg, 31(4), pp. 1253–1262. doi: 10.1007/s00468-017-1543-5.

749 Vogel, B. and Henstra, D. (2015) ‘Studying local climate adaptation: A heuristic research framework  
750 for comparative policy analysis’, *Global Environmental Change*. Elsevier Ltd, 31, pp. 110–120. doi:  
751 10.1016/j.gloenvcha.2015.01.001.

752 Vogt, J. *et al.* (2017) ‘Citree: A database supporting tree selection for urban areas in temperate  
753 climate’, *Landscape and Urban Planning*. Elsevier B.V., 157, pp. 14–25. doi:  
754 10.1016/j.landurbplan.2016.06.005.

755 Wang, S. *et al.* (2018) ‘Prediction of Lime Tolerance in Rhododendron Based on Herbarium  
756 Specimen and Geochemical Data’, *Frontiers in Plant Science*, 9(October), pp. 1–14. doi:

757 10.3389/fpls.2018.01538.

758 Wilson, A. (2016) *Future of Cities: The Science of Cities and Future Research Priorities*.

759 Government Office for Science.

760 Woodward, F. I., Lomas, M. R. and Kelly, C. K. (2004) ‘Global climate and the distribution of plant

761 biomes’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1450), pp.

762 1465–1476. doi: 10.1098/rstb.2004.1525.

763 Wright, I. J. *et al.* (2004) ‘The worldwide leaf economics spectrum’, *Nature*. doi:

764 10.1038/nature02403.

765 Wright, I. J. *et al.* (2017) ‘Global climatic drivers of leaf size’, *Science*, 357, pp. 917–921.

766 Yim, Y.-J. and Kira, T. (1975) ‘Distribution of forest vegetation and climate in the Korean peninsula’,

767 *Japanese Journal of Ecology*, 25(2), pp. 77–88.

768

769