DISPERSAL PROCESSES DRIVING PLANT MOVEMENT: RANGE SHIFTS IN A CHANGING WORLD

Pervasive human-mediated large-scale invasion: analysis of spread patterns and their underlying mechanisms in 17 of China's worst invasive plants

Nir Horvitz[1]*[†], Rui Wang[2][†], Fang-Hao Wan[2]* and Ran Nathan[1]

[1]*Department of Ecology, Evolution and Behavior, Movement Ecology Laboratory, Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel; and [2]State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China*

Summary

1. Biological invasions constitute a major component of human-induced environmental change and have become a world-wide problem threatening global biodiversity and incurring massive economic costs. Consequently, research on biological invasions proliferates, placing a major emphasis on species traits and habitat characteristics associated with successful invasion. Yet, the mechanisms underlying rapid spread and the resulting patterns remain largely unexplored.

2. Using data collected since 1980 and earlier at the county level all over China, we studied the contribution of potential dispersal vectors – railroads, rail stations, roads, general human activity, rivers and winds – to the spread of 17 of China's worst invasive plant species. Focusing on long-distance dispersal events, we calculated the minimal arrival speed for the first record of each species in each county. We also developed and applied a new method to account for observation bias due to the proximity to roads, using observational data of 776 native (non-invasive) plant species throughout China.

3. We found that human-related vectors are accountable for the vast spread of all 17 invasive plant species we examined. Spread patterns were characterized by long jumps of tens to hundreds of kilometres and extremely fast average spread rates of roughly 2–4 km per year, and a very broad range (0·1–128·2 km per year) with high variability between years. These rates are much higher than those expected from classic dispersal vectors such as water, wind or animals. Commonly used fat-tailed dispersal kernels did not fit the observed distribution of long jumps for any species.

4. *Synthesis*. We found pervasive empirical evidence for the overriding role of humans in the large-scale spread of invasive plants from multiple taxa. The observed spread patterns differ significantly from those portrayed in the literature, emphasizing the need to develop new frameworks to explore large-scale spread in general and invasive spread in particular. With public data sets of invasive species observations becoming increasingly more available, the time is ripe to go beyond exploration of species traits and habitat suitability and to examine the actual patterns and the mechanisms of large-scale invasive spread, even at a scale of thousands of kilometres over land.

Key-words: dispersal vectors, invasion ecology, invasive species, invasive spread, jump dispersal, long-distance dispersal, minimal arrival speed, plant dispersal, spread pattern

Introduction

Biological invasions – the entry, establishment and spread of non-native species – are considered a major cause of human-induced environmental change and have become a pressing problem in every biome on earth (Elton 1958; Williamson 1996; Pimentel, Zuniga & Morrison 2005; MacIsaac, Tedla & Ricciardi 2011). Biological invasions threaten global biodiversity by altering the structure and function of ecosystems (Levine *et al.* 2003; Trakhtenbrot *et al.* 2005; Cook *et al.* 2007) and disrupting key biological interactions (Mitchell *et al.* 2006; Traveset & Richardson 2006). These invasions have also been considered a major cause of recent extinctions (Thomas *et al.* 2004) and have had a substantial economic

*Correspondence authors. E-mails: nir.horvitz@mail.huji.ac.il, wanfanghao@caas.cn

[†]These authors contributed equally to this work

impact (Pimentel, Zuniga & Morrison 2005). The problem of biological invasions is expected to be exacerbated in light of global climate change, and several lines of evidence have predicted a rapid increase in the spread and damage caused by invasive species (Dukes & Mooney 1999; Walther *et al.* 2009; Bradley *et al.* 2010).

Invasive species, defined as non-indigenous species that have established and reached a widespread distribution in an area far beyond their native range (Colautti & MacIsaac 2004), undergo a process of spatial spread we termed here as 'invasive spread'. Similar to the spread of indigenous plant species in their native range, the process of invasive plant spread is shaped by three distinct, yet non-mutually exclusive, components: fecundity (the number of seeds produced and dispersed), dispersal kernel (the probability of a dispersed seed to arrive at a given location with respect to its mother plant) and habitat suitability (the probability of a seed to reach the reproductive stage). There has been extensive research on the roles of habitat suitability (Rouget & Richardson 2003; Morisette *et al.* 2006; Evangelista *et al.* 2008; Ibanez *et al.* 2014) and fecundity (Parker 2000; Lockwood, Cassey & Blackburn 2005; Mason *et al.* 2008) in determining invasive spread. Many studies have successfully quantified dispersal kernels of invasive plants over relatively short spatial scales (Buckley *et al.* 2005; Skarpaas & Shea 2007; Caplat, Nathan & Buckley 2012) or applied dispersal kernels to examine theoretical scenarios over large scales (Engler & Guisan 2009; Nathan *et al.* 2011a; Bullock *et al.* 2012). Yet, our ability to quantify dispersal kernels underlying actual large-scale spread of invasive species is still rather poor (but see Chapman *et al.* 2016), and the challenge of elucidating the relative contribution of different (and often unknown) dispersal vectors (Nathan *et al.* 2008) remains largely unresolved. Therefore, progress in understanding and predicting actual large-scale invasive spread still awaits empirically derived mechanistic models capable of representing spread over large areas (i.e. hundreds to thousands of kilometres). The challenge of quantifying dispersal kernels underlying an actual large-scale invasive spread is rather overwhelming, involving extremely rare but critical events that are inherently difficult to quantify and predict (Nathan 2006), multiple potential dispersal vectors and other interacting, and usually unknown, factors (Higgins, Nathan & Cain 2003b; Higgins *et al.* 2003a; Nathan 2006). Furthermore, the spread of invasive plants is almost always meagerly documented, and the available information typically appears in the form of sparse anecdotal, historical snapshots of the invasive process. It has been suggested that due to such difficulties and uncertainties, the process of large-scale spatial spread is practically unpredictable (Clark *et al.* 2001). Without downplaying these difficult fundamental challenges, understanding the underlying mechanisms and developing predictive frameworks for large-scale invasive spread are still of critical importance for both basic and applied research of biological invasions (Trakhtenbrot *et al.* 2005; Jongejans, Skarpaas & Shea 2008).

Advanced models based on reaction–diffusion equations have been commonly used by ecologists to predict asymptotic rates of invasion (Skellam 1951; Nathan *et al.* 2011a). Predicted rates of spread were found to match observed rates demonstrating the value of this approach (Okubo *et al.* 1989; Andow *et al.* 1990; Allen *et al.* 1991). Yet, these models are based on the assumption that the dispersal kernel of an individual is characterized by a continuous function with a mode relatively close (i.e. <100 m) to the source. On the other hand, observed patterns of invasive plants are often characterized by very long jumps (Kot, Lewis & van den Driessche 1996; Buchan & Padilla 1999). Such extreme long-distance dispersal (LDD) events are rather rare, and therefore problematic to quantify or even recognize (Nathan *et al.* 2003). Therefore, evaluating plant invasive spread requires empirical data at sufficiently large spatial (100s km) and temporal (over a few decades) scales.

In this study, we used a rare large-scale data set in order to characterize the main patterns and mechanisms that drive the rapid spread of 17 of China's worst invasive plant species (Table 1) across mainland China. Using data of first arrival to the county (i.e. the location of the first observation in each invaded county, Fig. 1), we aimed to compare the potential of some key LDD mechanisms in determining the rapid spread of all 17 plant species. Because current GIS data are available at good quality only for the last 3–4 decades, we limit our analysis to observations made after 1980, which include a total of 4698 'first arrival' observations (276·4 data points per species on average, see Fig. S1, Supporting Information).

To assess the spread rate based on this unique data set, we defined a metric called 'minimal arrival speed' (MAS, see Materials and methods), the minimal speed the species had to spread in order to get to each 'first arrival' observed location at the observed year. Minimal arrival speed values were compared to spread rates reported or expected for classic 'natural' (non-human) LDD mechanisms such as hydrochory, zoochory and anemochory. The former two mechanisms were assessed based on previously published studies (see Discussion); the latter was calculated as the theoretical spread rate expected by extreme winds. We then compared different human- (e.g. distance from roads, population) and habitat-related [distance from rivers, normalized difference vegetation index (NDVI)] distributions of potentially influential factors for each species invasion to those expected by random sampling from each invaded county. For example, to assess the contribution of cars to the spread of invasive plants through jump dispersal, we tested if the distances from the nearest road to the first arrival locations differ from those expected by random. We accounted for a sampling bias towards more accessible areas (Reddy & Dávalos 2003) by comparing our data set to the observational data of 776 non-invasive species (Table S1). Our null hypotheses were that (i) spread rates will match those reported for classic 'natural' dispersal factors, and that (ii) for each human- or habitat-related factor, observed distributions will not differ significantly from those expected by random. Our alternative hypothesis was that invasive plants spread mostly by long jumps mediated by humans that act as extreme LDD vectors. Hence, we alternatively predict that (i) MAS values would be much higher than those reported or

Table 1. Main traits of the study species. Main traits of the 17 study species. Abbreviations for known dispersal mechanisms are: WD – wind dispersed, WT – water dispersed, HM – dispersed by humans and AN – dispersed by animals. Introduction pathways are: I – human intentional, A – human accidental, N – natural, U – unknown

Species	Family	Longevity	Origin	Dispersal mechanism	Dispersal months	Invaded counties since 1980	Introduction pathways
Ageratina adenophora	Asteraceae	Perennial	Mexico	WD, WT	3–6	126	N
Ageratum conyzoides	Asteraceae	Annual, Perennial	S. America	WD, HM	1–12	295	I
Alternanthera philoxeroides	Amaranthaceae	Perennial	S. America	WT, HM	1–12	389	I
Amaranthus retroflexus	Amaranthaceae	Annual	USA	AN, WD	8–10	306	A
Amaranthus spinosus	Amaranthaceae	Annual	Trop. America	WD, WT, HM	7–11	328	A
Amaranthus viridis	Amaranthaceae	Annual	Trop. N. America	WD, WT, HM	8–10	356	A
Ambrosia artemisiifolia	Asteraceae	Annual	N. America	AN, WT, HM	9–11	385	A
Aster subulatus	Asteraceae	Annual, Biennial	N. America	WD, WT, HM	8–10	289	A
Chenopodium ambrosioides	Amaranthaceae	Annual	Mexico	AN, HM	4–1	356	U
Conyza canadensis	Asteraceae	Annual, Biennial	N. and C. America	WD, WT, HM	5–9	347	U
Erigeron annuus	Asteraceae	Annual	N. America	WD, WT, HM	1–9	380	U
Eupatorium odoratum	Asteraceae	Perennial	N. America	WD, HM	4–12	184	N
Flaveria bidentis	Asteraceae	Annual	S. America	AN	8–10	123	A
Galinsoga parviflora	Asteraceae	Annual	S. America	WD, HM, WT	7–10	307	U
Lepidium virginicum	Brassicaceae	Annual, Biennial, Perennial	N. and C. America	AN, HM	5–9	307	U
Phytolacca americana	Phytolaccaceae	Perennial	N. America	AN, HM	8–10	310	A
Solanum rostratum	Solanaceae	Annual	N. America	WT, AN, WD	8–10	47	A

calculated for the known natural dispersal vectors of the study species (Table 1), and (ii) that the distribution of human-related factors will differ significantly from the null expectation of random distribution.

Materials and methods

STUDY SPECIES AND DATA COLLECTION

China – a vast country that varies tremendously in topography, climate and other environmental features – has served as a continent-scale arena for the establishment and spread of many invasive alien species over the last decades. At least 268 invasive alien plant species are known in China, some of them, and especially the 17 species selected for this study, have caused severe damage to national economy and development and to local ecosystems and native species (Wan, Guo & Zhang 2009; Ma, Yan & Shou 2013). In turn, these invasive plant species have received considerable attention from local residents, scientists and the government, and substantial resources were allocated to control or manage them. As such, information about the invasive spread of these species is well documented, although it spans over multiple sources. We thus collected data from various available sources, including herbarium records, monitoring programmes operated by locals and by the central governments, scientific collections, published literature and our own field surveys, started in 2002 and is still ongoing. These large efforts yielded an exceptionally rich data set that portrays the invasion process of the study plants sufficiently well. The 17 study species were eventually selected as those presenting the greatest risk to China's economy and environment (Ma, Yan & Shou 2013), as well as those for which the first arrival year of presence was known in the first invaded county in China and for all subsequently invaded counties as well. Criteria for selecting the 17 species did not include invasion speed, human effects or number of invaded counties. Yet, the risk criterion restricted selected species to those that have invaded more than one natural area (Ma, Yan & Shou 2013), hence the selected species were relatively broadly

distributed. Overall, our data set includes 10 835 records from 1868, but for most analyses we focused on the 4698 first arrival records post 1980 (see *Relevant position index* below). The exact location of each record was available in most cases, and all records were faithfully assigned at the county level. China has 2391 counties (and 1415 counties in our data set) that are 63·2 ± 47·7 km apart (mean ± SD of the distance between central position of two neighbouring counties; range 0·6–670 km). Thus, analysing our data set at the county level allows a focus on LDD events at unusually large scales with exceptionally large sample sizes.

The 17 study species (Table 1) encompass five families and differ in longevity (nine annuals, two annuals or biennials, four perennials and two either annuals or perennials), origin (10 N. America, 5 S. America and 2 from Central America), height (five are 0·4–0·6 m tall, six are between 1·2–2 m tall and the rest between 0·7–1 m tall) and other important traits such as previously observed dispersal mechanisms, typical dispersal season and introduction pathway (for details, see Table 1).

MINIMAL ARRIVAL SPEED

The 'MAS' was defined as the minimal dispersal speed a species had to spread in order to get to a new specific location (i.e. newly invaded county) at a specific time (i.e. year of first observation at the county). For each earliest observed location in a county not previously invaded, we took the distance from all previously observed invaded points of the same species, divided those distances by the elapsed time for each observation and took the minimal speed of all calculated speeds as the species MAS at the specific location.

PROBABILITY DISTRIBUTION FITTING

Observed distributions of MAS values of each species were tested for goodness-of-fit to five commonly used probability distribution functions: Gaussian, Exponential, Weibull, Log-normal and WALD. Because our data were mostly collected at the county level
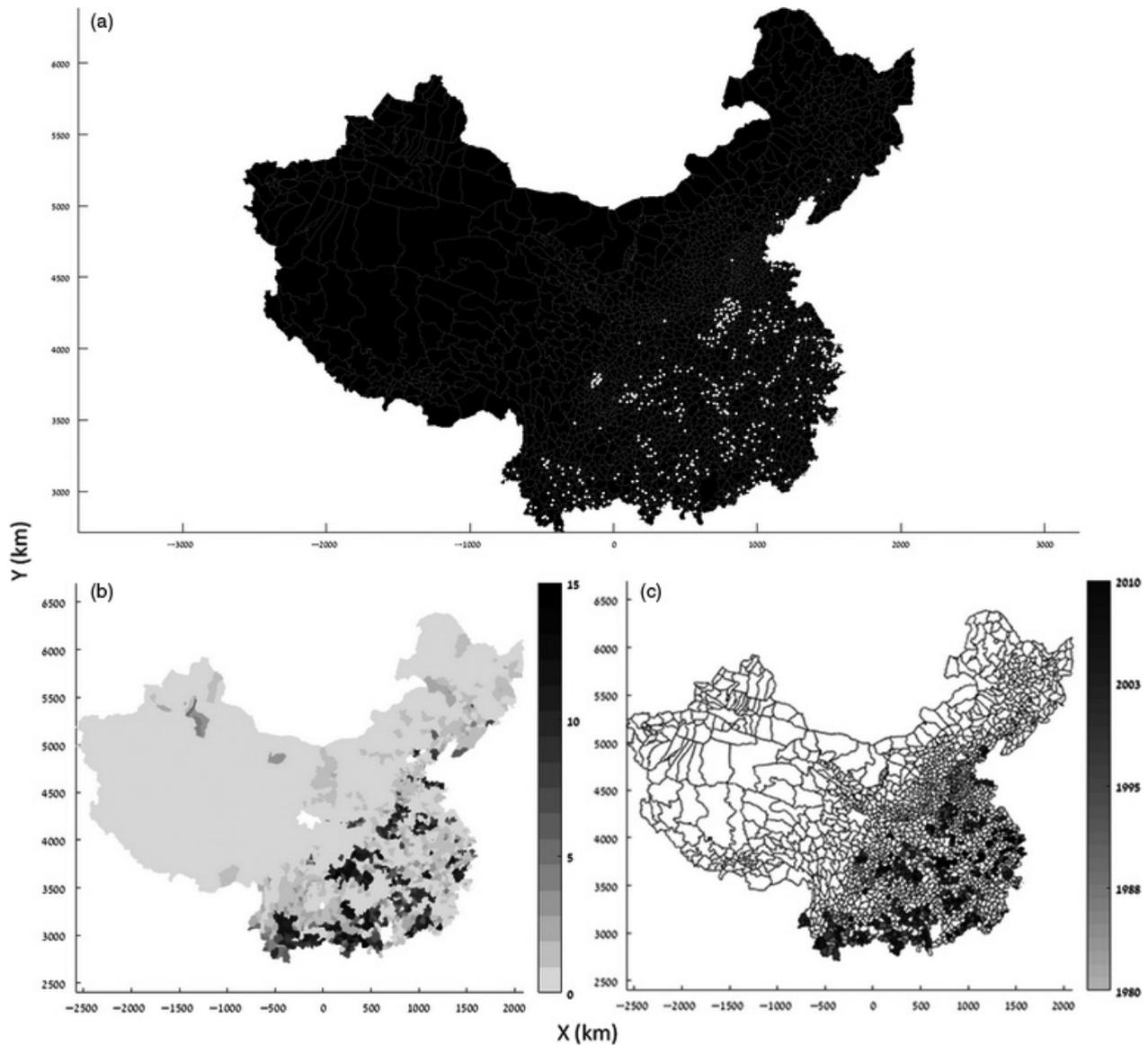
Fig. 1. Plant data used in the study. (a) Example of first arrival data. White dots are the locations of first observation of the plant species *Alternanthera philoxeroides* at the county level. (b) Total number of observations of invasive species at the county level. The data include only cases where the species invaded the county after 1980. (c) Example of the invasive spread of the plant species *A. philoxeroides*. Black polygons are counties where invasion was observed before 1980, white polygons are counties where the invasive species was not observed until 2010.

(repeated observations within the same county were much less documented), we focused on first arrival to a county (MAS) as explained above. This implies that calculated MAS cannot account for local (within-county) dispersal. We therefore fitted the data to truncated distributions, setting the minimal observed MAS value for each species (Table 2) as the lower bound threshold. The area under the curve beyond this threshold was normalized to unity for all five distributions.

RELEVANT POSITION INDEX OF POTENTIALLY INFLUENCING FACTOR IN A COUNTY

For each species in each invaded county, we looked at the different potentially influential factors (e.g. population, distance from road) at the first observation point in the county with respect to its distribution

throughout the county. Our null hypothesis is that the random variable representing the relevant position (ranging from 0 to 1) of a factor on the county distribution will be uniformly distributed. Thus, if for a certain species, the distribution of the relative position of a factor across all counties invaded by it will differ significantly from a uniform distribution, we conclude that this factor, or some other mechanism tightly correlated with this factor, should be considered in explaining the observed invasive spread. Because habitat suitability may affect the probability of each location to be the first arrival position, we reran our calculations while normalizing for habitat suitability and also compensated for sampling bias to accessible areas (see *Compensating for sampling bias*). Although some historical records were 148 years old, we limited our analysis to observations collected on 1980 or later, to match the period for which GIS data were available. GIS data on China county borders, distance from settlements

Table 2. Minimal arrival speed. The minimal arrival speed of the 17 study species. For all species, average MAS was over 1·5 km per year. Over 75% of the observations of each species were over 1 km per year, and for all but a single species, at least 10% of MAS was over 5 km per year

	Minimal arrival speed (km per year)					
Species	Average	SD	Min	Max	>1 (%)	>5 (%)
Ageratina adenophora	7·7	4·3	1·1	22·1	100	65
Ageratum conyzoides	2·4	2	0·2	9·9	75·8	10·9
Alternanthera philoxeroides	5·8	5	0·4	25·5	95·3	41
Amaranthus retroflexus	3·1	2·7	0·2	18·3	88·4	13·2
Amaranthus spinosus	2·8	2	0·2	10·2	82·5	11·7
Amaranthus viridis	2·3	1·7	0·2	9·5	76·9	10
Ambrosia artemisiifolia	4·8	4·7	0·2	52·3	92·2	31·6
Aster subulatus	5·3	3	0·1	17·2	96·5	47·2
Chenopodium ambrosioides	3·7	3·5	0·2	40	91·5	25·2
Conyza canadensis	1·6	1·1	0·1	6·5	66·9	1·5
Erigeron annuus	3·5	2·6	0·03	16·2	90·2	20·2
Eupatorium odoratum	3·8	2·3	0·3	14·1	95·1	22·5
Flaveria bidentis	9·9	8·6	1·2	66·2	100	73·1
Galinsoga parviflora	4·3	2·8	0·4	19·4	95	32·1
Lepidium virginicum	3·5	2·8	0·1	15·2	87·2	22·6
Phytolacca americana	3·6	2·9	0·3	32·1	92·6	18·4
Solanum rostratum	16·7	21	1·8	128·2	100	82·4
All species	3·9	4·2	0·1	128·2	88	24·7

(which was set as 0 inside settlement borders), distance from roads and distance from railways and rail stations were downloaded from CHGIS website (http://www.fas.harvard.edu/~chgis/data/dcw). Distance from rivers was calculated based on data downloaded from the HydroSHEDS site (http://hydrosheds.cr.usgs.gov). We defined a new 'major rivers' layer by taking only rivers with an 'Up_cells' attribute (a proxy for mean annual flow, see http://hydrosheds.cr.usgs.gov/hydro.php) higher than 2000. Data on human footprint was downloaded from 'The global footprint network' (http://www.footprintnetwork.org). All data were saved at 1 km × 1 km spatial resolution.

HABITAT SUITABILITY

Habitat suitability was modelled for each of the study species over the entire China mainland by the Maxent model (Phillips, Dudík & Schapire 2004), using all observational data (i.e. including pre-1980 and not first arrival to county data). Atmospheric input variables were based on the BioClim data set (including annual mean, minimum and maximum temperatures, annual and diurnal temperatures range, isothermality, annual precipitation and elevation), downloaded from the WorldClim web site (http://www.worldclim.org/bioclim). We used all atmospheric variables included in the data set. The habitat suitability raster for each species (see Fig. S2) was saved at 1 km × 1 km resolution.

COMPENSATING FOR SAMPLING BIAS

Observations of invasive species are expected to be biased towards locations more accessible by human such as roads (Kadmon, Farber & Danin 2004). In order to compensate for this bias, we ran the same analysis for 776 non-invasive plant species found in China (see Table S1). We then normalized our results for the position index of each factor by the results of the non-invasive species. The data on the 776 species was downloaded from GBIF (http://www.gbif.org/) and was limited to all non-invasive species with at least five observations in China and after 1980. We note that limiting the same analysis to species with at least 20 observations yielded similar results.

STATISTICAL ANALYSES

All statistical analyses were performed using Matlab version 7.14 (MathWorks Inc, Natick, MA, USA).

Results

For all study species, mean MAS (Table 2) was >1·5 km per year. At least 75% (and up to 100%) of the observed MAS of each species were >1 km per year. We did not find any spatial or directional structure in the results. Analysis of goodness-of-fit between MAS distribution of each species and key distribution functions (Gaussian, Exponential, Weibull, Log-normal and WALD) allowed for rejection of the null hypothesis of fit for all distribution functions and all species (Anderson-Darling test, Table S2). The frequency distribution of MAS values for all species combined has a unimodal shape, with a peak around 2·2 km per year (Fig. 2; see Fig. S3 for species-specific histograms).

To examine anemochory as a plausible dispersal mechanism, we used a model based on extreme statistics (Nathan *et al.* 2011a) to simulate the spread rate by extreme winds (see Table S3). Even under extreme average wind speeds of 30 m s^{-1}, which are over an order of magnitude higher than the maximum annual mean anywhere in China in all analysed years (Guo, Xu & Hu 2011), simulated spread speeds
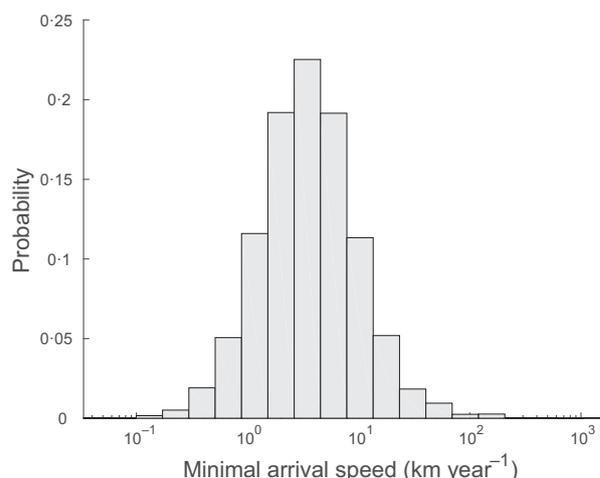

Fig. 2. Frequency distribution of minimal arrival speed for study species. A log-scaled histogram of the combined observed minimal arrival speed values of all 17 study species. The data range between 0·1 to 128·2 km per year, with a peak around 2·2 km per year.

were at least one order of magnitude lower than the observed speeds.

Comparing the spatial distribution of the observations against those of potentially influential factors (Fig. 3), weighted by habitat suitability and with compensation for sampling bias, we found significant bias towards human-related factors (distances from rail stations, roads and settlements, human footprint index and population size) in all species except *A. adenophora* and *S. rostratum*. All species but the aforementioned two were also strongly biased ($P < 0.001$) towards low NDVI values. We found no significant bias towards short distances from all rivers or only from major rivers in any of the species. We note that if not using a statistical correction for a number of tests, a significant trend towards short distances from rivers is found in 15 of the 17 species. We ran the analyses without weighting for habitat suitability and found similar results but with somewhat higher significance levels; we also found similar results without compensating for sampling bias, but the significance levels were much higher. We found no effect of the introduction pathway on MAS distribution.

To further examine the strong trend towards human-related factors, we also assessed the correlation between the number of private cars in possession of Chinese citizens (National Bureau of Statistics of China, http://www.stats.gov.cn) and the number of counties with invasive plants observations (Fig. 4). The logarithmic model $y = a + b\log(x)$ fitted the data well ($R^2 = 0.96$, $P < 0.001$ for all species combined; R^2 ranged from 0.91 to 0.99 for each species separately, each with $P < 0.001$).

Discussion

Seventeen of China's worst invasive plants spread over the country extremely quickly (Table 2). For all species, at least 75% of the 'new arrival' observations suggested average minimal speeds of over 1 km per year, with the most extreme rates ranging between 6·5 (*Conyza canadensis*) and 128·2 (*Solanum rostratum*) km per year. Such spread rates are more than an order of magnitude faster than those predicted by any of the study species suggested dispersal mechanisms, such as wind (see Table S3), animals (Viana *et al.* 2013), rivers (Boedeltje *et al.* 2003) or even cars (Von der Lippe & Kowarik 2007; von der Lippe *et al.* 2013). Theoretically, dispersal by migrating animals, most likely birds, could explain such vast spread (Viana *et al.* 2016). Yet, given the lack of directionality in spread pattern and the high frequency of LDD events in all study species, we find it unlikely that zoochory, even by migratory birds, plays a major role in the spread processes of the study species. Contrasting the lack of support for classic 'native' dispersal vectors such as animals, water or wind, we found pervasive evidence for strong effects of human activity. For nearly all study species, first observations in a county were significantly closer to potentially influencing factors related to human-mediated dispersal mechanisms compared to the null random expectations (Fig. 3). These first observations were characterized by short distances from roads and settlements, high population density, high human footprint index and proximity to rail stations, even after compensating for the sampling bias towards human accessible areas. Similar conclusions were drawn for the fast invasive spread of the Pinewood nematode (*Bursaphelenchus xylophilus*) throughout China (Robinet *et al.* 2009), suggesting that human-mediated invasion might be generalized across different life-forms and taxonomic groups.

First observations were also characterized by low NDVI values. We suggest two possible explanations for this finding. First, regions with frequent human activity, with many roads and settlements, are usually characterized by low NDVI values. Thus, the low NDVI values might indirectly reflect human-mediated dispersal as well. Alternatively, most study species are known to inhabit a broad array of habitats (Liu *et al.* 2005; Qiu, Shalimu & Tan 2013), including low fertility soils and several other habitats that are unsuitable to most other plants. Furthermore, this result might reflect higher establishment success in areas with lower NDVI due to reduced competition. Previous studies showed that invasive plants are more likely to establish in regions with both high (Gavier-Pizarro *et al.* 2010; Huang *et al.* 2012) or low NDVI (Dukes & Mooney 1999), although such variation might be attributed to the mismatch between the coarser spatiotemporal scale of analysis and the relevant finer scale required for assessing NDVI effects. This mismatch also occurs in our study, impeding valid distinction between these two possible explanations. More generally, habitat associations are probably better investigated through field experiments rather than by analysing large-scale observational data, but this is beyond the scope of this work.

The evidence we found for pervasive, human-mediated spread of invasive plants corresponds well with the emerging consensus on invasive spread (Nathan 2006; Nathan *et al.* 2008; Wilson *et al.* 2009). Our study strongly supports this view due to the exceptionally large spatial scale, long time-period and multiple species included in the analysis. These merits, however, inherently hinder attempts to address the ensuing question of which specific human-related factors are responsible for this rapid spread. This important question requires a data set that can break down the multicollinearity of the highly correlated explanatory (human-related) variables. This challenge might be accomplished, for example, by examining trends in a spatiotemporal resolution sufficient to tease apart different human-related factors. This is inherently difficult to achieve at the large spatiotemporal scales covered in our study, whereas studies at smaller scales might lead to different results reflecting specific features of a smaller area. For example, our previous analysis of the spread rate of *Ageratina adenophora* (Horvitz *et al.* 2014) revealed that rivers accounted for much faster spread than roads, in contrast to the current study results (Fig. 3). The area analysed in the previous study was less than 8% of the area included in the current study, and within this area, traffic was concentrated on major roads that follow rivers, thus the effects of the two vectors were difficult to tease apart. Furthermore, even the strongest correlation we found, between the number of
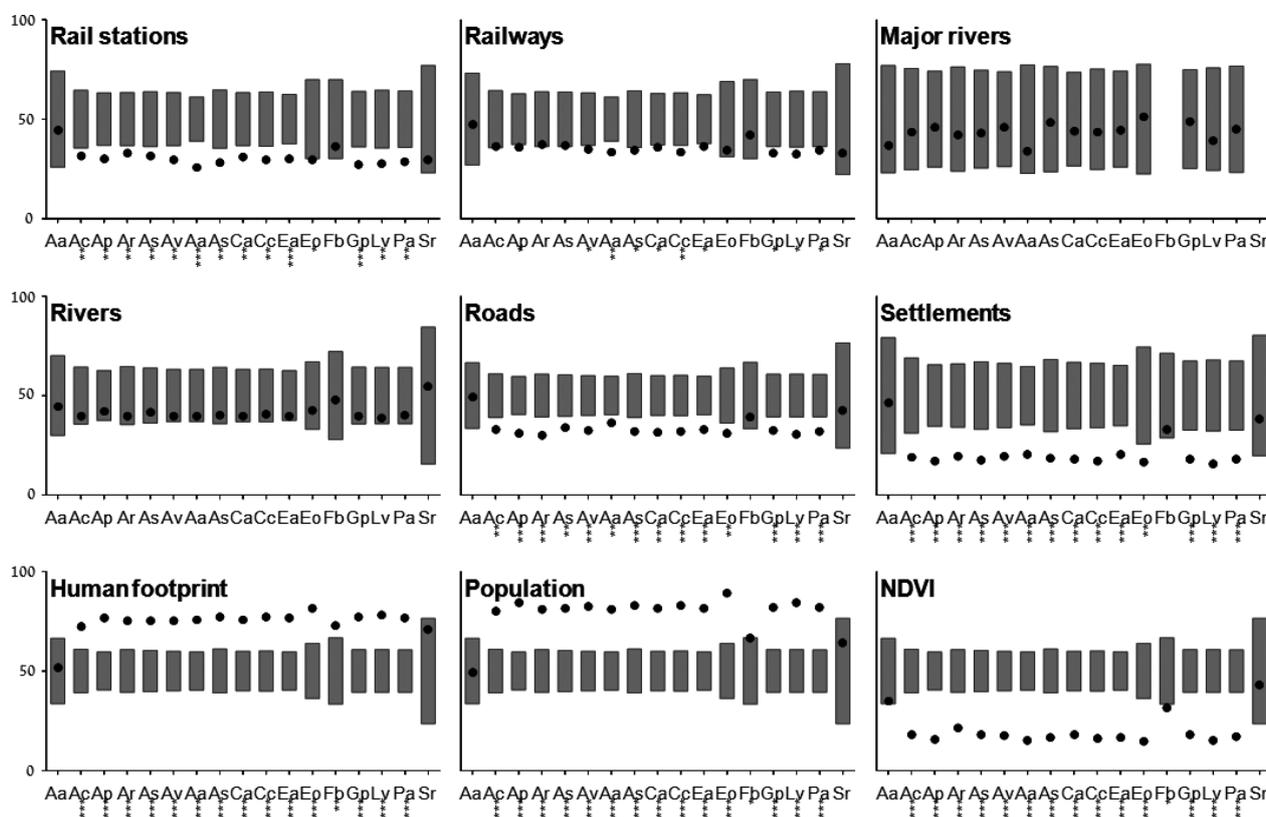
Fig. 3. Results of comparison between first arrival location and county distribution of several factors. Results of mean distribution of the 17 study species in comparison to uniform (random) distribution of nine environmental features. For each county, each factor was checked to see where it falls on its distribution throughout the county. Data were weighted by Maxent model results and observation bias was compensated for using the observations of non-invasive plant species distributions. Black dots represent the mean species value. Grey bars represent the 95% confidence bounds expected from a uniform distribution. One, two or three asterisks under species name represent critical significance levels of 0·05, 0·01 and 0·001 respectively. Columns in 'major rivers' panel for species with less than five observations in counties containing a major river are left empty.

counties with observations of invasive species and the number of private cars in possession of Chinese citizens (Fig. 4), should not imply that cars are necessarily the main dispersal vector of invasive plants in China. Other factors related to China's fast economical growth, such as urbanization or freight traffic, are also expected to show an increase similar to car possession. Altogether, these factors reflect the ultra-fast urbanization process of China (Xu *et al.* 2016), suggesting that large urban areas – with large volume of traffic going in and out – might act as a key vertex in the spread process of invasive plants, similar to the role of airports in epidemic spread (Colizza *et al.* 2006).

The importance of LDD has already been highlighted in early essays on dispersal and biogeography (Darwin 1859; Higgins & Richardson 1999). Kot, Lewis & van den Driessche (1996) showed that classic reaction–diffusion models using exponentially bounded (e.g. Gaussian) dispersal kernels can drastically underestimate observed rates of spread, whereas fat-tailed kernels (e.g. log-normal) adequately account for fast and even accelerating spread (see also Higgins & Richardson 1999). In fat-tailed dispersal kernels, the very low probability of dispersal to long distances decays slowly with increasing distance, and LDD events can be perceived as stochastic draws of (the rare) large values from the (fat) distribution tail. This implies tremendous variation in the

dispersal distances of individual LDD events, or in other words, invasive spread through big and highly variable jumps, extending over 4 orders of magnitude (Fig. 2). This pattern was anticipated theoretically (Nathan 2006) but has rarely been documented in previous studies of human-mediated LDD, presumably because empirical studies are inherently limited to (relatively) short distances and short periods (Wich-mann *et al.* 2009; Taylor *et al.* 2012; von der Lippe *et al.* 2013). Furthermore, most theoretical works either highlighted potential LDD mechanisms (Nathan *et al.* 2008; Wilson *et al.* 2009) and general spread patterns (Nathan *et al.* 2008; Wilson *et al.* 2009) or presented detailed models of spread patterns (Lewis 1997) or particular dispersal mechanisms (e.g. Nathan *et al.* 2011b). Yet, they become excessively complicated and difficult to validate when attempting to link patterns and mechanisms and to account for multiple dispersal vectors (Nathan 2007). Common to most theoretical and empirical studies is the basic assumption portraying invasive spread as a spatially continuous process, characterized by the (very) low probabilities of extreme LDD (or jump dispersal) events at the dispersal kernel tail (Nathan *et al.* 2012). This basic assumption prevails in both classic (Skellam 1951; Okubo *et al.* 1989; Allen *et al.* 1991) and contemporary literature (Adams *et al.* 2015; Uden *et al.* 2015; Teller, Zhang & Shea 2016). However, the findings of the current study suggest that
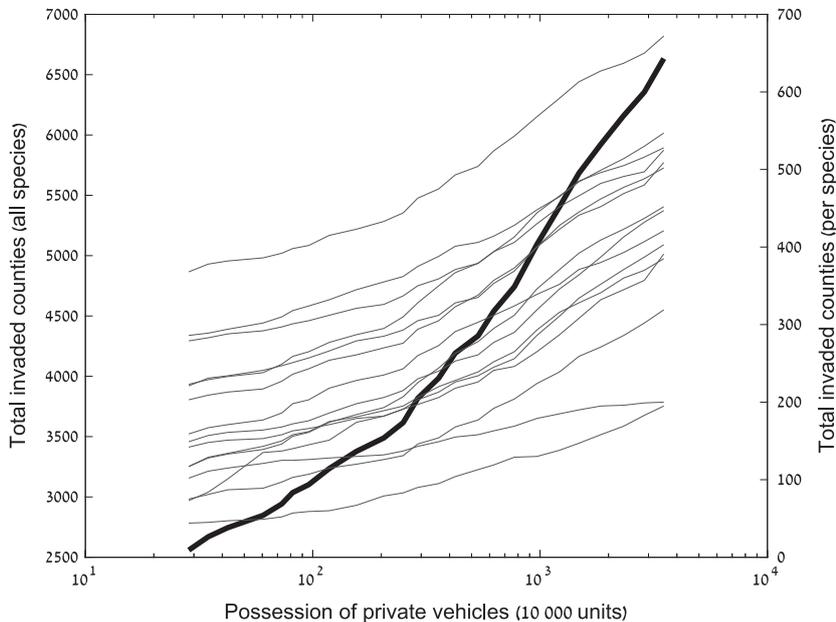
Fig. 4. Relationship between the spread of invasive plants and the number of cars in a county. Invasive spread was estimated as the number of counties where invasive plant species were observed. The number of private cars possessed by Chinese citizens is the explanatory variable. Narrow lines represent the number of invaded counties per species (right ordinate). Thick lines represent total number of counties for all species (left ordinate). Best fit to the logarithmic model $y = a + b\log(x)$ is $R^2 = 0.96$, $P < 0.001$. Data encompass years 1985–2010.

the occurrence of such long jumps cannot be consistently estimated based on a continuous fat-tailed dispersal kernel.

The exceptionally large scope of our data set and analysis provided a unique opportunity to rigorously test the assumption that fat-tailed dispersal kernels can properly represent invasive spread driven by jump dispersal. Such a pattern is evident in all 17 study species, which have rapidly expanded their range within China by means of very long jumps of highly variable lengths. However, the distribution of MAS did not fit any of five probability density functions (dispersal kernels) commonly used in the literature, including two exponentially bounded and three fat-tailed ones (Table S2). Because our analysis was (intentionally) biased to the tail of the distribution – by examining first arrival to a new county which is on average tens of kilometres from the nearest neighbouring one (see Materials and methods) – we fitted the data to truncated distributions restricted to values equal to or higher than the minimal MAS found for each species. Although the lack of fit to exponentially bounded kernels was expected, we found that fat-tailed kernels poorly fit the data as well. Hence, standard continuous distribution functions commonly used in invasive spread models, including fat-tailed ones that can well represent local establishment or invasive spread in small areas such as islands, cannot properly depict the stochastic human-mediated long jumps that characterize the rapid invasive spread of plants.

In this study, we presented a new approach to study the mechanisms responsible for rapid invasive spread of plants, a widespread phenomenon observed world-wide over the last decades (Pyšek & Hulme 2005). Our study provides pervasive evidence that invasive plants can spread very rapidly, and that human-mediated jump dispersal is the only reasonable explanation for this phenomenon. We recommend further use of MAS or similar metrics to explore rapid invasive spread and advocate controlling for observational bias towards human accessible areas, as well as the use of the consistently increasing, free online data sets in invasion biology research. Our study is one of the very few using empirical data sets covering very large scales to investigate invasive spread patterns (Chapman *et al.* 2016). The finding that commonly used fat-tailed dispersal kernels cannot describe and predict invasive spread mediated by human-mediated jump dispersal does not necessarily imply that such patterns are inherently unpredictable and their underlying mechanisms cannot be understood (Clark *et al.* 2001). Subsequent efforts should be allocated to elucidate the specific human activities responsible for rapid invasive spread of plants in China and elsewhere, providing the means to guide management plans and mitigate the damage inflicted by invasive species to global biodiversity and economy.

Authors' contributions

R.W. and F.-H.W. collected the data; N.H., R.W. and R.N. analysed the data. All authors contributed to the writing of the manuscript.

Acknowledgements

Data accessibility

GIS data on China county borders, distance from settlements distance from roads and distance from railways and rail stations were downloaded from CHGIS website (http://www.fas.harvard.edu/~chgis/data).

Distance from rivers was calculated based on data downloaded from the HydroSHEDS site (http://hydrosheds.cr.usgs.gov).

Data on human footprint were downloaded from 'The global footprint network' (http://www.footprintnetwork.org/en/index.php/GFN/page/footprint_data_and_results).

Environmental data are based on the BioClim data set available on the WorldClim web site (http://www.worldclim.org/bioclim) and are included in the Supporting Information.

Species data were downloaded from GBIF (http://www.gbif.org/) and are included in the Supporting Information.

Data about Chinese citizens were downloaded from National Bureau of Statistics of China (http://www.stats.gov.cn/english/).

References

Adams, V.M., Petty, A.M., Douglas, M.M., Buckley, Y.M., Ferdinands, K.B., Okazaki, T., Ko, D.W. & Setterfield, S.A. (2015) Distribution, demography and dispersal model of spatial spread of invasive plant populations with limited data. *Methods in Ecology and Evolution*, **6**, 782–794.

Allen, L.J.S., Allen, E.J., Kunst, C.R.G. & Sosebee, R.E. (1991) A diffusion-model for dispersal of Opuntia-Imbricata (Cholla) on rangeland. *Journal of Ecology*, **79**, 1123–1135.

Andow, D.A., Kareiva, P.M., Levin, S. & Okubo, A. (1990) Spread of invading organisms. *Landscape Ecology*, **4**, 177–188.

Boedeltje, G., Bakker, J.P., Bekker, R.M., Van Groenendael, J.M. & Soesbergen, M. (2003) Plant dispersal in a lowland stream in relation to occurrence and three specific life-history traits of the species in the species pool. *Journal of Ecology*, **91**, 855–866.

Bradley, B.A., Blumenthal, D.M., Wilcove, D.S. & Ziska, L.H. (2010) Predicting plant invasions in an era of global change. *Trends in Ecology & Evolution*, **25**, 310–318.

Buchan, L.A. & Padilla, D.K. (1999) Estimating the probability of long-distance overland dispersal of invading aquatic species. *Ecological Applications*, **9**, 254–265.

Buckley, Y.M., Brockerhoff, E., Langer, L., Ledgard, N., North, H. & Rees, M. (2005) Slowing down a pine invasion despite uncertainty in demography and dispersal. *Journal of Applied Ecology*, **42**, 1020–1030.

Bullock, J.M., White, S.M., Prudhomme, C., Tansey, C., Perea, R. & Hooftman, D.A. (2012) Modelling spread of British wind-dispersed plants under future wind speeds in a changing climate. *Journal of Ecology*, **100**, 104–115.

Caplat, P., Nathan, R. & Buckley, Y.M. (2012) Seed terminal velocity, wind turbulence, and demography drive the spread of an invasive tree in an analytical model. *Ecology*, **93**, 368–377.

Chapman, D.S., Makra, L., Albertini, R., Bonini, M., Paldy, A., Rodinkova, V., Sikoparija, B., Weryszko-Chmielewska, E. & Bullock, J.M. (2016) Modelling the introduction and spread of non-native species: international trade and climate change drive ragweed invasion. *Global Change Biology*, in press.

Clark, J.S., Carpenter, S.R., Barber, M. *et al.* (2001) Ecological forecasts: an emerging imperative. *Science*, **293**, 657–660.

Colautti, R.I. & MacIsaac, H.J. (2004) A neutral terminology to define 'invasive' species. *Diversity and Distributions*, **10**, 135–141.

Colizza, V., Barrat, A., Barthélemy, M. & Vespignani, A. (2006) The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 2015–2020.

Cook, D.C., Thomas, M.B., Cunningham, S.A., Anderson, D.L. & De Barro, P.J. (2007) Predicting the economic impact of an invasive species on an ecosystem service. *Ecological Applications*, **17**, 1832–1840.

Darwin, C. (1859) *On the Origins of Species by Means of Natural Selection.* Murray, London, UK.

Dukes, J.S. & Mooney, H.A. (1999) Does global change increase the success of biological invaders? *Trends in Ecology & Evolution*, **14**, 135–139.

Elton, C.S. (1958) *The Ecology of Invasions by Animals and Plants.* Methuen, London, UK.

Engler, R. & Guisan, A. (2009) MigClim: predicting plant distribution and dispersal in a changing climate. *Diversity and Distributions*, **15**, 590–601.

Evangelista, P.H., Kumar, S., Stohlgren, T.J., Jarnevich, C.S., Crall, A.W., Norman, J.B. III & Barnett, D.T. (2008) Modelling invasion for a habitat generalist and a specialist plant species. *Diversity and Distributions*, **14**, 808–817.

Gavier-Pizarro, G.I., Radeloff, V.C., Stewart, S.I., Huebner, C.D. & Keuler, N.S. (2010) Housing is positively associated with invasive exotic plant species richness in New England, USA. *Ecological Applications*, **20**, 1913–1925.

Guo, H., Xu, M. & Hu, Q. (2011) Changes in near-surface wind speed in China: 1969–2005. *International Journal of Climatology*, **31**, 349–358.

Higgins, S.I., Nathan, R. & Cain, M.L. (2003b) Are long-distance dispersal events in plants usually caused by nonstandard means of dispersal? *Ecology*, **84**, 1945–1956.

Higgins, S.I. & Richardson, D.M. (1999) Predicting plant migration rates in a changing world: the role of long-distance dispersal. *American Naturalist*, **153**, 464–475.

Higgins, S.I., Clark, J.S., Nathan, R., Hovestadt, T., Schurr, F., Fragoso, J.M.V., Aguiar, M.R., Ribbens, E. & Lavorel, S. (2003a) Forecasting plant migration rates: managing uncertainty for risk assessment. *Journal of Ecology*, **91**, 341–347.

Horvitz, N., Wang, R., Zhu, M., Wan, F.H. & Nathan, R. (2014) A simple modeling approach to elucidate the main transport processes and predict invasive spread: river-mediated invasion of *Ageratina adenophora* in China. *Water Resources Research*, **50**, 9738–9747.

Huang, D., Zhang, R., Kim, K.C. & Suarez, A.V. (2012) Spatial pattern and determinants of the first detection locations of invasive alien species in mainland China. *PLoS ONE*, **7**, e31734.

Ibanez, I., Diez, J.M., Miller, L.P. *et al.* (2014) Integrated assessment of biological invasions. *Ecological Applications*, **24**, 25–37.

Jongejans, E., Skarpaas, O. & Shea, K. (2008) Dispersal, demography and spatial population models for conservation and control management. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 153–170.

Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, **14**, 401–413.

Kot, M., Lewis, M.A. & van den Driessche, P. (1996) Dispersal data and the spread of invading organisms. *Ecology*, **77**, 2027–2042.

Levine, J.M., Vila, M., Antonio, C.M., Dukes, J.S., Grigulis, K. & Lavorel, S. (2003) Mechanisms underlying the impacts of exotic plant invasions. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 775–781.

Lewis, M.A. (1997) Variability, patchiness, and jump dispersal in the spread of an invading population. *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions* (eds D. Tilman & P.M. Kareiva), pp. 46–69. Princeton University Press, Princeton, NJ, USA.

von der Lippe, M., Bullock, J.M., Kowarik, I., Knopp, T. & Wichmann, M. (2013) Human-mediated dispersal of seeds by the airflow of vehicles. *PLoS ONE*, **8**, e52733.

Liu, J., Liang, S.C., Liu, F.H., Wang, R.Q. & Dong, M. (2005) Invasive alien plant species in China: regional distribution patterns. *Diversity and Distributions*, **11**, 341–347.

Lockwood, J.L., Cassey, P. & Blackburn, T. (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, **20**, 223–228.

Ma, J., Yan, X. & Shou, H. (2013) *The Checklist of the Chinese Invasive Plants.* Higher Education Press, Beijing, China.

MacIsaac, H.J., Tedla, R.A. & Ricciardi, A. (2011) Patterns and rate of growth of studies in invasion ecology. *Fifty Years of Invasion Ecology: The Legacy of Charles Elton* (ed. D.M. Richardson), pp. 51–60. Wiley-Blackwell, New York, NY, USA.

Mason, R.A., Cooke, J., Moles, A.T. & Leishman, M.R. (2008) Reproductive output of invasive versus native plants. *Global Ecology and Biogeography*, **17**, 633–640.

Mitchell, C.E., Agrawal, A.A., Bever, J.D. *et al.* (2006) Biotic interactions and plant invasions. *Ecology Letters*, **9**, 726–740.

Morisette, J.T., Jarnevich, C.S., Ullah, A., Cai, W., Pedelty, J.A., Gentle, J.E., Stohlgren, T.J. & Schnase, J.L. (2006) A tamarisk habitat suitability map for the continental United States. *Frontiers in Ecology and the Environment*, **4**, 11–17.

Nathan, R. (2006) Long-distance dispersal of plants. *Science*, **313**, 786–788.

Nathan, R. (2007) Total dispersal kernels and the evaluation of diversity and similarity in complex dispersal systems. *Seed Dispersal: Theory and Its Application in a Changing World* (ed. A.J. Dennis), pp. 252–276. CABI, Wallingford, UK.

Nathan, R., Perry, G., Cronin, J.T., Strand, A.E. & Cain, M.L. (2003) Methods for estimating long-distance dispersal. *Oikos*, **103**, 261–273.

Nathan, R., Schurr, F.M., Spiegel, O., Steinitz, O., Trakhtenbrot, A. & Tsoar, A. (2008) Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution*, **23**, 638–647.

Nathan, R., Horvitz, N., He, Y., Kuparinen, A., Schurr, F.M. & Katul, G.G. (2011a) Spread of North American wind-dispersed trees in future environments. *Ecology Letters*, **14**, 211–219.

Nathan, R., Katul, G.G., Bohrer, G., Kuparinen, A., Soons, M.B., Thompson, S.E., Trakhtenbrot, A. & Horn, H.S. (2011b) Mechanistic models of seed dispersal by wind. *Theoretical Ecology*, **4**, 113–132.

Nathan, R., Klein, E., Robledo-Arnuncio, J.J. & Revilla, E. (2012) Dispersal kernels: review. *Dispersal Ecology and Evolution* (eds J. Clobert, M. Baguette, T.G. Benton & J.M. Bullock), pp. 187–210. Oxford University Press, Oxford, UK.

Okubo, A., Maini, P.K., Williamson, M.H. & Murray, J.D. (1989) On the spatial spread of the Grey squirrel in Britain. *Proceedings of the Royal Society of London B: Biological Sciences*, **238**, 113–125.

Parker, I.M. (2000) Invasion dynamics of *Cytisus scoparius*: a matrix model approach. *Ecological Applications*, **10**, 726–743.

Phillips, S.J., Dudík, M. & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *Proceedings of the Twenty-First International Conference on Machine Learning* (ed. C. Brodley), pp. 83. ACM, Banff, Canada.

Pimentel, D., Zuniga, R. & Morrison, D. (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, **52**, 273–288.

Pyšek, P. & Hulme, P.E. (2005) Spatio-temporal dynamics of plant invasions: linking pattern to process. *Ecoscience*, **12**, 302–315.

Qiu, J., Shalimu, D. & Tan, D.Y. (2013) Reproductive characteristics of the invasive species *Solanum rostratum* in different habitats of Xinjiang, China. *Biodiversity Science*, **21**, 590–600.

Reddy, S. & Dávalos, L.M. (2003) Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, **30**, 1719–1727.

Robinet, C., Roques, A., Pan, H., Fang, G., Ye, J., Zhang, Y. & Sun, J. (2009) Role of human-mediated dispersal in the spread of the pinewood nematode in China. *PLoS ONE*, **4**, e4646.

Rouget, M. & Richardson, D.M. (2003) Inferring process from pattern in plant invasions: a semimechanistic model incorporating propagule pressure and environmental factors. *The American Naturalist*, **162**, 713–724.

Skarpaas, O. & Shea, K. (2007) Dispersal patterns, dispersal mechanisms, and invasion wave speeds for invasive thistles. *The American Naturalist*, **170**, 421–430.

Skellam, J.G. (1951) Random dispersal in theoretical populations. *Biometrika*, **38**, 196–218.

Taylor, K., Brummer, T., Taper, M.L., Wing, A. & Rew, L.J. (2012) Human-mediated long-distance dispersal: an empirical evaluation of seed dispersal by vehicles. *Diversity and Distributions*, **18**, 942–951.

Teller, B.J., Zhang, R. & Shea, K. (2016) Seed release in a changing climate: initiation of movement increases spread of an invasive species under simulated climate warming. *Diversity and Distributions*, **22**, 708–716.

Thomas, C.D., Cameron, A., Green, R.E. *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.

Trakhtenbrot, A., Nathan, R., Perry, G. & Richardson, D.M. (2005) The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions*, **11**, 173–181.

Traveset, A. & Richardson, D.M. (2006) Biological invasions as disruptors of plant reproductive mutualisms. *Trends in Ecology & Evolution*, **21**, 208–216.

Uden, D.R., Allen, C.R., Angeler, D.G., Corral, L. & Fricke, K.A. (2015) Adaptive invasive species distribution models: a framework for modeling incipient invasions. *Biological Invasions*, **17**, 2831–2850.

Viana, D.S., Santamaría, L., Michot, T.C. & Figuerola, J. (2013) Allometric scaling of long-distance seed dispersal by migratory birds. *The American Naturalist*, **181**, 649–662.

Viana, D.S., Gangoso, L., Bouten, W. & Figuerola, J. (2016) Overseas seed dispersal by migratory birds. *Proceedings of the Royal Society of London B: Biological Sciences*, **283**, 20150426.

Von der Lippe, M. & Kowarik, I. (2007) Long-distance dispersal of plants by vehicles as a driver of plant invasions. *Conservation Biology*, **21**, 986–996.

Walther, G.-R., Roques, A., Hulme, P.E. *et al.* (2009) Alien species in a warmer world: risks and opportunities. *Trends in Ecology & Evolution*, **24**, 686–693.

Wan, F.H., Guo, J.Y. & Zhang, F. (2009) *Research on Biological Invasions in China*. Science Press, Beijing, China.

Wichmann, M.C., Alexander, M.J., Soons, M.B. *et al.* (2009) Human-mediated dispersal of seeds over long distances. *Proceedings of the Royal Society of London B: Biological Sciences*, **276**, 523–532.

Williamson, M. (1996) *Biological Invasions*. Chapman & Hall, London, UK.

Wilson, J.R., Dormontt, E.E., Prentis, P.J., Lowe, A.J. & Richardson, D.M. (2009) Something in the way you move: dispersal pathways affect invasion success. *Trends in Ecology & Evolution*, **24**, 136–144.

Xu, M., He, C., Liu, Z. & Dou, Y. (2016) How did urban land expand in China between 1992 and 2015? A multi-scale landscape analysis *PLoS ONE*, **11**, e0154839.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Historical spread of the 17 study species throughout China.

Fig. S2. Maxent model predictions of habitat suitability for the study species.

Fig. S3. Frequency distribution of minimal arrival speed.

Table S1. Observations of non-invasive plants species in mainland China.

Table S2. Possible distribution functions of mean arrival speed.

Table S3. Observed spread rate compared to wind-dispersal theoretical dispersal distances.