Latitudinal gradients in taxonomic overdescription rate affect 2 macroecological inferences using species list data 3 4 Owen R. Jones^{1,2,*}, Andy Purvis^{1,3}, and Donald L.J. Quicke^{1,3,4} 5 6 7 1 Division of Biology, Imperial College London, Silwood Park Campus, Ascot, Berkshire, 8 SL5 7PY, UK 9 2 Current address: Max Planck Institute for Demographic Research, Konrad Zuse Str 1, 10 Rostock, D-18057, Germany 11 3 Centre for Population Biology, Imperial College London, Silwood Park Campus, Ascot, 12 Berkshire, SL5 7PY, UK 13 4 Department of Entomology, Natural History Museum, Cromwell Road, London SW7 5BD, 14 UK 15 16 17 *Corresponding author: jones@demogr.mpg.de, telephone: +493812081125; fax: +493812081425 18 Running header: Latitudinal patterns in taxonomic overdescription 19 20

1 Abstract

2 Species lists for particular geographic areas are often used in macroecology and conservation; 3 for example, they have been used to identify hotspots of biological diversity, and for the 4 study of latitudinal species diversity gradients. However, there is uncertainty over the 5 accuracy of species lists due to undiscovered species and synonymy of described species. 6 This paper concentrates on taxonomic overdescription caused by the latter. Where bias in the 7 amount of taxonomic overdescription occurs along a variable of interest (e.g. latitude, or 8 body size), inferences from macroecological studies may be affected. This form of taxonomic 9 overdescription is likely to be most pronounced in speciose groups, where many species have 10 relatively small geographic range sizes and have low numerical abundance. A good example 11 of such a group is the wasp family Ichneumonidae. We first use taxonomic and region-12 specific species list data for the Ichneumonidae to estimate the probability of species validity. Then we use this estimated probability to statistically correct the region's species richness 13 estimate using a Monte Carlo simulation approach, and examine the effect the correction has 14 on three major macroecological patterns: the relative species richness of geographic regions, 15 16 latitudinal species richness pattern, and body size. Our results indicate that although there is significant geographic variation in overdescription, the bias is not sufficient to qualitatively 17 18 alter broad-scale macroecological conclusions such as hotspot identity, and the qualitative global patterns of diversity and mean body size. 19

1 Introduction

2 In macroecology and hotspot-based conservation, the fundamental unit is usually the species 3 (Magurran 2004). Species checklists -- lists of species that are common to a particular geographic area -- have been put to use in these areas. For example, they have been used to 4 5 prioritise conservation funding (Myers et al. 2000), for monitoring communities (Roberts 6 et al. 2007), and studying latitudinal gradients in diversity, geographic range size, and body 7 size (Rosenzweig 1995, Gaston & Blackburn 2000). It follows, therefore, that bias in species 8 lists will have implications for research in these areas. The (usually unwritten) assumption for 9 most macroecological studies is that the concept of species is the same across all of the units 10 in the study and that there is no systematic bias relative to the hypothesis. If it were not for 11 the uneven nature of these biases we could consider species uncertainty as random error. 12 However, the facts that some taxonomic groups are better known than others, some regions are better known than others, and both body size and range size have been identified as 13 14 correlates of the probability of description (Blackburn and Gaston 1995, Collen et al. 2004, Jones et al. 2009) indicate that the error is probably not random. For example, in a study of 15 16 latitudinal biodiversity gradients, workers typically assume equal confidence in the identification of species, and in the proportion of species known, along the whole latitudinal 17 18 range of the study. If this assumption is not met then our confidence in, for example, a 19 gradient showing higher diversity in the tropics might be shaken: it could be that the increased diversity is simply because a greater proportion of the species have been identified 20 21 there, or a greater amount of splitting (raising subspecies to species level) has been carried 22 out there.

1 The nature of species lists

2 Current species checklists are a result of accumulated research through time. They are not 3 static; they change as new species are discovered and as systematists conduct revisionary 4 work (Alroy 2002). Since species are used as data points in macroecology, the number of 5 data points in an analysis, and their spatial distribution, will change with time. The rate of 6 discovery of new species is directly related to the amount of applied sampling effort; thus 7 areas where research is more intense will appear to have more species than areas with less 8 research, other things being equal. Sampling effort, and its effect on apparent species 9 richness, has received attention in the literature and several statistical approaches exist to deal 10 with it (McCabe and Gotelli 2000, Gotelli, 2001, Gotelli and Colwell, 2001). Differences in 11 taxonomic treatment between geographic regions may also introduce bias and is cited as a 12 serious problem for orchid conservation (Pillon and Chase 2007).

13 Changes in species concept with time, and taxonomic error

14 The particular species concept used for an analysis is likely to have a large effect, not only on 15 the number of species in a given area but also on the location of areas of relatively high 16 species richness and endemism (Townsend Peterson 2006) -- an obvious problem in conservation planning. An explicit species concept is not usually adopted in publications but 17 18 it is clear that these concepts have changed through time. The biological species concept has 19 been widely accepted since the 1940s and it is only relatively recently that there has been a 20 trend away from the biological species concept towards a finer-grained phylogenetic species 21 concept (Wheeler and Meier 2000, Agapow et al. 2004, Isaac et al. 2004), in part because it is 22 easy to use with phylogenetic data (Sites and Marshall 2004). In addition, the dominant 23 species concept differs among taxonomic groups, even for those to which the same set of

concepts are potentially applicable. For example, in the mid-twentieth century there was a
 move among ant taxonomists to eradicate infraspecific names (Wilson and Brown 1953),
 while other taxonomists (e.g. butterfly taxonomists) embraced infraspecific names by
 adopting a polytypic species concept (Mayr 1963). Thus, counts of species from different
 taxonomic groups may not be comparable.

6 Allied to changes in species concepts are the personal approaches of individual systematists 7 who encounter new specimens and undertake taxonomic revisions. Their view can be 8 regarded as falling along a spectrum from "lumper" to "splitter" (Quicke 1993). Lumpers 9 emphasise the similarities rather than the differences between species, and are less likely to 10 establish new species and more likely to establish synonymies by merging existing species. 11 Splitters, on the other hand, are more likely to emphasise differences between species and are 12 thus more likely to establish new species and less likely to establish synonymies. An overabundance of splitters can lead to overdescription. In other words, some species are 13 14 thought to be valid but are, in fact, merely synonyms of a single species. This issue has been discussed for plants of the Indian sub-continent (Venu 2002), and has also been regarded as a 15 16 problem with hotspot identification for amphibians (Chaitra et al. 2004). The related problem of taxonomic inflation caused by the artificial elevation of subspecies to the species level 17 18 (Isaac et al. 2004) is not considered here.

Taxonomic revision of genera can reduce the number of species by lumping two or more species into a single species (synonymisation). Such revisions are not carried out at random. In insects they are usually undertaken at a regional scale and on a relatively small taxonomic group - often a genus (Jones et al. 2009). Revisions can establish new species by description of new material or by splitting an existing species into two or more species. The distribution of the spectrum of lumper/splitter approaches among taxonomic groups and through space

1 and time is likely to vary, introducing yet more bias into species lists. This bias is 2 compounded by the similarly uneven distribution of revisionary taxonomy among taxa, 3 regions and through time. A further problem is that revisions take time, and are relatively 4 infrequent compared to the rate of description of new species. Thus, the more time that 5 elapses since the original description of a species, the greater the probability that a revision 6 will be undertaken to "test" whether it is a valid species or merely a synonym (Alroy 2002, 7 Alroy 2003). Of course, multiple revisions may be carried out on some taxa and species may 8 switch from synonym to valid species and back again. Nevertheless, unless the prevailing 9 species concept changes, it is probable that older taxa are more stable than younger taxa. 10 Lastly, the revision process relies heavily on a body of knowledge about the genus (or other 11 taxonomic group) and about the region from which it is derived. As a taxonomic group and 12 its regional setting become better known, taxonomic revisions carried out on the group 13 become more likely to reach the correct conclusion about species placement. Thus, the 14 probability of a correct designation (as valid or synonym) is likely to be affected by the diversity of the group and region, which together influence how well the region is known. 15 16 Previous studies of taxonomic inexactness have concentrated on groups where the great 17 majority of species are already known; usually terrestrial vertebrates. The vast majority of animal species are insects and, coincidentally, this group is also believed to be one of the 18 19 least completely known. Therefore, it is appropriate to investigate the effects of description bias in a group of relatively poorly known insects. Ichneumonidae is an ideal group for this 20 21 because they are geographically widespread and, although the group has been studied for a 22 considerable time, it is clear from species accumulation curves, and the proportion of species 23 that are known from only few specimens, that they are still relatively poorly known.

1 We use a published taxonomic dataset on the Ichneumonidae (Yu et al. 2005) to investigate 2 the effects of systematic bias in the amount of overdescription on the inferences that we can 3 make from species list data. The Ichneumonidae are a classic example of where the almost universal observation of increased species richness in the tropics compared to temperate 4 5 zones is not followed (Sime and Brower 1998). Although some authors have cited differences 6 in taxonomic effort as the main cause (e.g. Morrison et al. 1979), repeated surveys have 7 produced the same puzzling result (Hawkins 1994). Numerous theories have been put 8 forward to explain this contrary observation including resource fragmentation (Janzen 1981), 9 intensity of predation (Rathcke and Price 1976), and levels of chemical defence (Gauld et al. 10 1992): see Sime and Brower (1998) for an overview. We hypothesise that latidudinal 11 variation in synonymy rates may exist and that this may affect species richness estimates and 12 the outcome of analyses where latitude is a covariate. To address this we first examine the 13 taxonomic record to determine the existence of spatial or temporal bias in synonymy rate. We 14 consider spatial bias to exist when species from different geographic areas are treated 15 differently. For example, workers in some regions may be more likely to lump species 16 together, creating synonyms, than in others. Temporal bias can exist when such taxonomic 17 treatment has changed through time. For example, if there has been a transition from a 18 splitting approach to a lumping approach. We then turn to macroecology to examine the 19 effects of these biases on the relative species richness of geographic regions, and latitudinal 20 gradients in species richness and body size. Our analysis allows us to estimate the effect of 21 taxonomic treatment on macroecological patterns at both a regional and global scale.

1 Methods

2 Ichneumonid biological data

The family, genus, species, authority, year of description and status (i.e. valid taxon, synonym etc.) for each species (or synonym) in the Ichneumonidae were extracted from the Taxapad database of Yu et al. (2005). For most of the species, body length information was also available. Where more than one body length measurement was available for a species, the maximum was taken and, where no measure for a particular species was available, the mean for the genus was used.

9 Geographic information

10 Geographic location data were also extracted from the database. The majority (88%) of 11 species within the database are associated with a list of countries (or regions within country) 12 where they have been recorded. This data allowed us to assign species records to continent, 13 region and subregion, using a scheme developed by the Taxonomic Diversity Working Group 14 (TDWG) (Brummitt 2001). A summary of this scheme is provided in supplementary material (Appendix 1, Table A1). The continental regions we considered were: Africa, Europe, 15 Northern America, Asia Temperate, Asia Tropical, Australasia, Pacific and Southern 16 America (Appendix 1, Figure A1). We also obtained data on latitude and longitude (taken to 17 be the mid-points of the recorded locality) and geographic area (km²) for each TDWG region 18 19 and subregion. In the TDWG scheme, Central America (from the southern border of Mexico to the Panama-Colombia border) was included in the Southern America region (Figure A1, 20 21 and maps in Brummitt 2001).

22 Modelling taxonomic overdescription

23 We modelled the probability that a specimen was currently valid as a function of its body

1	length, the time elapsed since discovery, the total geographic area of the reported locations (a
2	measure of geographic range size), and the continental region (as defined by the TDWG in
3	Brummitt 2001). To carry out the modelling we used a generalised linear model (GLM), with
4	binomial errors and logit link. We included each term as a main effect and we also fitted
5	body size, time elapsed since discovery, and geographic range size in two-way interactions
6	with continental region. We scored validity as 1 (presently viewed as valid) or 0 (presently
7	viewed as non-valid). Non-valid species included synonyms, nomen nuda, and species
8	recorded as having an "unknown status".
9	We selected these terms because they are known to be important correlates of the taxonomic
10	record. For example, body size, description date, and geographic range size are reported to
11	correlate with probability of description in some groups (Blackburn and Gaston, 1995, Collen
12	et al. 2004), and species concepts have changed through time (Wheeler and Meier 2000,
13	Agapow et al. 2004, Isaac et al. 2004). In addition, we include geographic location because is
14	also possible that workers in different geographic areas take dissimilar taxonomic
15	approaches.
16	For each species we summed the spatial area of the subregions where each species had been
17	recorded to provide an estimate of range size. We then assigned each species to the continent
18	where the majority of its range fell. Again, we used the TDWG's geographic definitions for
19	this (Brummitt 2001 and Appendix 1, Table A1). For example, if a species occurred in
20	regions belonging to Europe and Africa, but 60% of its range was in Europe, it was regarded

as European.

Because of the very large number of data points (22,081), we set the alpha value for our
modelling to 0.001 to avoid effects too weak to have any biological meaning. To allow us to

1 validate the fit of the model we bootstrapped the model with 1000 iterations. On each 2 iteration, we randomly selected (with replacement) 90% of the data to fit the GLM model 3 described above and reserved 10% of the data for validation. To determine the optimum 4 probability threshold at which to assign records to validity classes (0=non-valid, 1=valid) we 5 calculated the Receiver Operating Characteristic (ROC) of the model (Fielding and Bell 6 1997). The ROC tests a range of probability thresholds at which a prediction is assigned to 7 being positive (valid) or negative (non-valid). It then compares the predicted values to the 8 actual values and calculates the proportions of true and false positives. The optimum value 9 for the threshold is the point at which the proportion of true positives is maximized and false 10 positives is minimized.

To validate the model, on each of the bootstrap iterations, we calculated Cohen's Kappa statistic (Cohen 1968), a measure of classification accuracy for categorical items, using the independent subset of the data (i.e. the 10% of the data that were not used to train the model). In addition, we calculated the area under the ROC curve, which provides another estimate of the ability of the model to correctly classify items (Fielding and Bell 1997).

16

17 Macroecological patterns

The model described above provided us with a fitted probability of validity for each species. We used this fitted value in a Monte Carlo simulation approach to generate 1000 statistically corrected taxonomies as follows: For each species we drew a random number (a 0 or 1) from a binomial distribution with a probability of success (drawing a 1) provided by the model's fitted value (i.e. the probability-of-validity) for the species in question. We repeated this procedure for the entire dataset 1000 times to produce a distribution of 1000 corrected

1	taxonomies. Then, using each of these 1000 taxonomies, we examined three macroecological
2	patterns; (1) relative regional species richness; (2) latitudinal species richness gradient; and
3	(3) latitudinal body size gradient.

- 4
- 5

Relative species richness of geographic regions

Assessments of regional species richness have been used to identify biodiversity hotspots
(Myers et al. 2000). We therefore examined the species richness of TDWG regions and
subregions (Brummitt 2001, and Appendix 1, Table A2) before and after statistical correction
for overdescription to see if correcting for overdescription affected the ranked richness
patterns of these areas.

We first calculated the number of species currently regarded to be valid in each region, and ranked the regions in order of species richness. We then repeated this calculation for each of the 1000 corrected taxonomies that are described above to produce 1000 species richness ranking estimates for each region, from which we could calculate a mean rank, and 95% confidence intervals for the estimate of the mean. We then plotted the predicted ranked richness with the current ranked richness, and examined the relationship with an ordinary least squares regression model.

18 We repeated this exercise at the TDWG geographic subregion level.

19

Latitudinal gradients: species richness, synonymy rate, and body size

20 Most species in the dataset (88%) were associated with at least one locality that had a known

- 21 latitude and longitude. We capitalized on this to examine latitudinal trends in species
- richness, synonymy rate, and body size, by dividing the world into 10° latitudinal bins and

examining the spatial variation in species richness, synonymy rate, and mean body size
 across the bins.

We first did this for currently valid species and then repeated the exercise using the 1000 statistically corrected taxonomies generated from the Monte Carlo simulations that we describe above. In this way we could examine how overdescription affects these latitudinal trends.

In addition, we investigated the effect of examining specific longitudinal regions in an
attempt to allow for the potential effect of the uneven distribution of effort, oceans and arid
zones. We did this by repeating the analyses with subsets of the data restricted to three broad
regions: the Americas (-165° to -30°), Europe and Africa (-15° to 60°) and an eastern region
(60° to 180°). We also repeated the analysis at smaller (5°) and larger (20°) latitudinal bin
sizes to check whether our results were consistent at different spatial scales.

13 **Results**

14 Taxomomic overdescription

15 Despite the small alpha value (0.001), the minimum adequate model included all of the 16 starting terms (Table 1). The cross-validation area under the true positive ROC curve was 17 0.886 (95% confidence interval (CI) < 0.001), and the Kappa statistic was 0.523 (95% CI = 18 0.001) (from the 1000 bootstrapped iterations). Landis and Koch (1977) would regard this to be a moderate agreement value and, because Kappa statistics often underestimate model 19 20 accuracy (Boyce et al. 2002), we consider our model to be reliable. The coefficient estimates for most continental-scale regions were fairly precise, but for the Pacific region, the 95% CIs 21 22 were large, indicating a high degree of uncertainty for this area.

1 The effects of bias on macroecological inferences

2 Relative species richness of geographic regions

The rankings of regions, and subregions, by species richness changed when the corrections were applied (Fig. 1, and Appendix 1, Tables A3 & A4). However, at both spatial scales, the relationship between mean ranking of the 1000 iterations of the Monte Carlo simulation, and the current rankings indicate that the overall trend was maintained, with the identity of the richest, and lowest, diversity areas remaining fairly consistent (Fig. 1 A-B). At both spatial scales, the relationship was almost exactly 1:1 with the slope of the fitted model not significantly different from 1 (Region scale: slope= 0.973 (95% CI = 0.066), t₄₈=-0.815,

10 p=0.419; subregion scale: slope= 0.99 (95% CI = 0.01), t_{282} =-1.041, p=0.299).

11 Variation in species richness, synonymy and body size

12 In the southern hemisphere there was a clear latitudinal gradient in species richness, with 13 species richness increasing towards the equator (Fig. 2A and Appendix 1, Table A2). In the 14 northern hemisphere the pattern was quite different. Here, the species richness pattern was 15 modal with greatest species richness found at 35° to 45° latitude (Fig. 2A). Applying the 16 statistical correction in the southern hemisphere reduced the gradient, but did not change the 17 qualitative pattern. Likewise, in the northern hemisphere, applying the statistical correction 18 for overdescription flattened the humped relationship, but did not remove the modality 19 entirely. These observations were broadly consistent across different longitudinal windows, which all showed that maximum species richness in the northern hemisphere occurred at 20 middle latitudes (35-55° N) (Appendix 1, Fig. A2A-C). 21

There was a gradient in synonym rate in the southern hemisphere, with rates declining
towards the equator (Fig. 2B). In the northern hemisphere, with the exception of the window

at 5-15° where there was a pronounced richness peak, there was no strong latitudinal trend 1 2 (Fig. 2B). These global-scale patterns were qualitatively consistent at both smaller (5°) and 3 larger (20°) latitudinal bin sizes (Appendix 1, Fig. A3). An examination of the synonymy patterns from the three longitudinal windows (Appendix 1, Fig. A2D-F), showed that the 4 5 global-scale peak in synonymy in the northern hemisphere was driven by high synonymy 6 rates in the Americas at the latitude of Central America (Appendix 1, Fig. A2D). In Europe 7 and Africa (-15° to 60°: Appendix 1, Fig. A2E) synonymy rates were slightly higher in the 8 northern hemisphere than in the southern hemisphere and peaked at 25° to 35° (north 9 Africa/Arabian Peninsula). In the eastern region (Appendix 1, Fig. A2F), synonymy rates 10 were highest in the southern hemisphere, and there was a gradient in synonymy in the 11 northern hemisphere where synonymy rates increased polewards. Thus it appears that the 12 global pattern may largely be driven by the patterns at eastern longitudes (60° to 180°). 13 Although, on a global scale, synonymy rates in the southern hemisphere appeared to be 14 negatively correlated with species richness, this pattern was not repeated in the northern 15 hemisphere, nor in the longitudinal subsets. 16 Mean body size varied modally with latitude, peaking at 5° to 15° of latitude globally (Fig. 17 2C). This pattern was broadly similar across the three longitudinal windows (Fig. 2 G-I), and

18 was consistent at different spatial scales (Appendix 1, Fig. A3E-F). However, in the

19 Americas (Appendix 1, Fig. A2H) body size peaked south of the equator at -25° to 15°

20 latitude, while in the Europe/Africa window (Appendix 1, Fig. A2I) it peaked in the northern

21 hemisphere at 15° to 25°, and in the eastern window (Appendix 1, Fig. A2J) it peaked at the

equator $(-5^{\circ} \text{ to } 5^{\circ})$. In all cases, the statistical correction for overdescription resulted in only a

23 slight reduction in mean body size.

1 Discussion

The proportion of ichneumonid species known from only 1 or 2 specimens is fairly large (34% from a sample of revision literature (O.R. Jones, unpublished data)). It is, therefore, not surprising that there is considerable uncertainty surrounding the correct designation of a species as valid or not valid. Such uncertainty is a feature common to many taxonomic groups and is likely to be most common in groups where species richness is high but where the numerical abundance of individuals within a species is low, and where geographic rangesize is small.

9 The apparent species richness patterns of ichneumonids were not qualitatively affected by our 10 correction for overdescription. After correction, species richness still increased towards the 11 equator in the southern hemisphere, and was still modal, with a peak at around 35° to 45°, in 12 the northern hemisphere. This pattern was consistent in each of the longitudinal windows we 13 examined, including the Americas, where our results in the northern hemisphere show 14 remarkably similar patterns to those observed by Janzen (1981).

15 In the southern hemisphere we found a striking trend for synonymy rate to increase towards 16 the pole, while in the northern hemisphere this trend was less clear, and there was a peak in synonymy rate at the latitude of Costa Rica. This probably reflects the high taxonomic flux 17 18 (Alroy 2002) that is inevitable when a species rich area like Costa Rica is subjected to 19 enormous taxonomic effort. The hymenoptera of Costa Rica have been intensively studied 20 since the mid-1980s when I.D. Gauld and D.H. Janzen established a program of Malaise trap surveys that has collected and examined literally millions of specimens (Gaston et al. 1996). 21 There was no clear positive relationship between species richness and synonymy rate. For 22 23 example, although in the global-scale analysis in the southern hemisphere, there was a

negative association between species richness and synonymy rate, the same pattern was not
apparent north of the equator. It was also not apparent in the three longitudinal windows. We
expected *a priori* that synonymy would be highest where species richness was highest,
reflecting the difficulty of working in a hyper-diverse region but it seems that this is not the
case.

6 Like Collen et al. (2004) and Blackburn and Gaston (1995), we found that the taxonomic record was correlated with a variety of factors. We found that probability of validity was 7 8 correlated with body size, range size, and geographic location, as well as the length of time 9 that the species has been established. Where these factors are random with respect to the 10 variables being studied, they can simply be regarded as random error and will pose no 11 statistical challenge. If, however, any of these factors vary in a systematic way along an axis 12 that is being studied, a gradient in the amount of taxonomic overdescription may be 13 generated. Such a gradient has the potential to undermine the results obtained from such a 14 study. In this study, we only corrected for overdescription and our methods did not attempt to correct for bias caused by differences in the discovery rate of new species and variation in 15 16 taxonomic effort (Isaac et al. 2004). However, taxonomic effort, in the sense of revisionary work, would clearly have an impact on synonymisation and therefore overdescription. This is 17 18 evidenced by the high synonymy rate in Costa Rica, an area that has received a lot of 19 taxonomic attention (Gaston et al. 1996)

Our methods provide a test, and method of partial correction, for overdescription and are applicable to any system where taxonomic data and adequate ancillary data are available for the species in the taxonomy. We expect that corrections for taxonomic overdescription will be required for relatively underworked groups such as most insect groups whose taxonomic databases are likely to be biased (Santos et al. 2010a), but may not be required for extremely

well-known groups such as terrestrial vertebrates. In addition, because the corrections should
only be required where there is systematic bias along an axis of interest, where the error is
evenly distributed the correction should not be necessary. This situation, for example, could
occur where the area studied is a relatively small geographic area that has been relatively
evenly studied by random sampling.

6 Our analysis of relative species richness of distinct geographic areas indicates that, even 7 though the validity of a large proportion of the ichneumonid species is brought into question, 8 the large-scale patterns may well be sufficiently strong to be qualitatively unaffected by such 9 correction. However, this is case-specific and it is easy to envisage cases where the 10 taxonomic overdescription biases could affect apparent hotspot identity (e.g. Townsend 11 Peterson 2006) or species area relationships (e.g. Santos et al. 2010b). For example, where 12 some regions have undergone substantial collection and primary description work, but little 13 revisionary work, apparent species richness is likely to be grossly inflated. We recommend 14 testing, and correcting, for bias on a case-by-case basis. The end-users of species lists and of the studies that use lists should be made aware of the potential problems inherent in their use. 15 16 Species lists are not static and are vulnerable to change because of new species being discovered and because of the instability in the species level nomenclature that is brought 17 18 about by revisionary taxonomy. Changes in taxonomic practices through time, may also be an 19 issue for some groups. Changes in species lists are inevitable and these changes have the potential to affect both the quantitative and qualitative outcomes of analyses which use them. 20

21

22 Acknowledgments

23 This work was supported by grant NE/C519583 from the UK Natural Environment Research

1	Council. We thank Natalie Cooper, Dalia Conde Ovando, Fernando Colchero, and three
2	anonymous referees for helpful comments on an earlier draft of this paper.
3	
4	References
5	Agapow, P. et al. 2004. The impact of species concept on biodiversity studies Quarterly

6 Review of Biology 79: 161-179.

Alroy, J. 2002. How many named species are valid? - Proceedings of the National Academy
of Sciences of the United States of America 99: 3706-3711.

9 Alroy, J. 2003. Taxonomic inflation and body mass distributions in North American fossil

10 mammals. - Journal of Mammalogy 84: 431-443.

11 Blackburn, T. and Gaston, K. 1995. What determines the probability of discovering a species

12 - a study of South-American oscine passerine birds. - Journal of Biogeography 22: 7-14.

Boyce, M.S. et al. 2002. Evaluating resource selection functions. - Ecological Modelling 157:
281-300.

15 Brummitt, R. 2001. World Geographical Scheme for Recording Plant Distributions. - Hunt

16 Institute for Botanical Documentation, Carnegie Mellon University, Pittsburgh.

17 Chaitra, M. et al. 2004. The biodiversity bandwagon: the splitters have it. - Current Science
18 86: 897-899.

19 Cohen, J. 1968. Weighted Kappa: nominal scale agreement with provision for scaled

20 disagreement or partial credit. - Psychological Bulletin 70: 213-220.

1	Collen, B. et al. 2004. Biological correlates of description date in carnivores and primates
2	Global Ecology and Biogeography 13: 459-467.
3	Fielding, A.H. and Bell, J.F. 1997. A review of methods for the assessment of prediction
4	errors in conservation presence/absence models Environmental Conservation 24: 38-49.
5	Gaston, K.J. and Blackburn, T.M. 2000. Pattern and process in macroecology Blackwell.
6	Gaston, K.J. et al. 1996. The size and composition of the hymenopteran fauna of Costa Rica.
7	- Journal of Biogeography 23: 105-113.
8	Gauld, I.D. et al. 1992. Plant allelochemicals, tritrophic interactions and the anomalous
9	diversity of tropical parasitoids: the 'nasty' host hypothesis Oikos 65: 353-357.
10	Gotelli, N.J. and Colwell, R.K. 2001. Quantifying biodiversity: procedures and pitfalls in the
11	measurement and comparison of species richness Ecology Letters 4: 379-391.
12	Gotelli, N.J. 2001. Research frontiers in null model analysis Global Ecology and
13	Biogeography 10: 337-343.
14	Hawkins, B.A. 1994. Pattern and process in host-parasitoid interactions Cambridge
15	University Press.
16	Isaac, N. et al. 2004. Taxonomic inflation: its influence on macroecology and conservation
17	Trends In Ecology and Evolution 19: 464-469.
18	Janzen, D. 1981. The peak in North American Ichneumonid species richness lies between 38°
19	and 42° N Ecology 62: 532-537.

20 Jones, O.R. et al. 2009. Using taxonomic revision data to estimate the geographic and

- 1 taxonomic distribution of undescribed species richness in the Braconidae (Hymenoptera:
- 2 Ichneumonoidea). Insect Conservation and Diversity 2: 204-212.
- 3 Landis, J.R. and Koch, G.G. 1977. The measurement of observer agreement for categorical
- 4 data. Biometrics 33: 159-319.
- 5 Magurran, A.E. 2004. Measuring biological diversity. Blackwell.
- 6 Mayr, E. 1963. Animal species and evolution. Harvard University Press.
- 7 McCabe, D.J. and Gotelli, N.J. (2000) Effects of disturbance frequency, intensity, and area on
- 8 assemblages of stream macroinvertebrates. Oecologia 124: 270-279.
- 9 Morrison, G. et al. 1979. Anomalous diversity of tropical parasitoids: A general
- 10 phenomenon? American Naturalist 114: 303-307.
- Myers, N. et al. 2000. Biodiversity hotspots for conservation priorities. Nature 403: 853858.
- 13 Pillon, Y. and Chase, M.W. 2007. Taxonomic exaggeration and its effects on orchid
- 14 conservation. Conservation Biology 21: 263-265.
- Quicke, D.L.J. 1993. Principles and Techniques of Contemporary Taxonomy. Chapman and
 Hall.
- Rathcke, B.J. and Price, P.W. 1976. Anomalous diversity of tropical icheumonid parasitoids:
 a predation hypothesis. American Naturalist 110: 889-893.
- 19 Roberts, R.L. et al. 2007. Using simple species lists to monitor trends in animal populations:
- 20 new methods and a comparison with independent data. Animal Conservation 10: 332-339.

1	Rosenzweig, M.L. 1995. Species diversity in space and time Cambridge University Press.
2	Santos, A.M.C. et al. 2010a. Assessing the reliability of biodiversity databases: identifying
3	evenly inventoried island parasitoid faunas (Hymenoptera: Ichneumonoidea) worldwide
4	Insect Conservation and Diversity 3: 72-82.
5	Santos, A.M.C. et al. 2010b. Are species-area relationships from entire archipelagos
6	congruent with those of their constituent islands? - Global Ecology and Biogeography 19:
7	527-540.
8	Sime, K.R. and Brower, A.V.Z. 1998. Explaining the latitudinal gradient anomaly in
9	ichneumonid species richness: evidence from butterflies Journal of Animal Ecology 67:
10	387-399.
11	Sites, J. and Marshall, J. 2004. Operational criteria for delimiting species Annual Review
12	of Ecology Evolution and Systematics 35: 199-227.
13	Townsend Peterson, A. 2006. Taxonomy is important in conservation: a preliminary
14	reassessment of Philippine species-level bird taxonomy Bird Conservation International 16
15	155-173.
16	Venu, P. 2002 Some conceptual and practical issues in taxonomic research Current Science
17	82,924-933.
18	Wheeler, Q. and Meier, R. 2000. The phylogenetic species concept (sensu Wheeler and
19	Platnick) In: Wheeler, Q. and Meier, R. (eds), Species concepts and phylogenetic theory: a
20	debate Columbia University Press, pp. 55-69.
21	Wilson, E. and Brown, W. 1953. The subspecies concept and its taxonomic application

- 1 Systematic Zoology 2: 97-111.
- 2 Yu, D. et al. 2005. World Ichneumonoidea 2004. Taxonomy, Biology, Morphology and
- 3 Distribution. CD/DVD. Taxapad, Vancouver, Canada.
- 4
- 5 Supplementary material (available as Appendix EXXXXX at
- 6 <www.oikosoffice.lu.se/appendix>). Appendix 1.
- 7

1 Figure headings

Figure 1: The relationship between current regional or subregional species richness rankings and the rankings predicted by a model accounting for taxonomic overdescription at (A) a regional scale and (B) a subregional scale. The relationship at the continental scale remained unchanged. Each point represents a region or subregion. The broken black line (partially obscured) is unity, while the grey unbroken line is the fit of a linear model of current and predicted ranking. Rankings and species richness data are provided in Tables A3 and A4 in Supplementary Material Appendix 1.

9

Figure 2: (A) Global species richness patterns in the Ichneumonidae. The solid black line 10 indicates the species richness as it is currently understood. The broken line indicates our 11 predictions for current diversity after accounting for taxonomic overdescription. The range of 12 13 these predictions is indicated by the shaded area (Table A2). (B) Proportion taxonomic 14 overdescription in the Ichneumonidae as a function of latitude. The line represents the mean 15 predicted synonymy, after correcting for overdescription, while the shaded area indicates the 16 range of the predictions. (C) Species body size as a function of latitude. The line represents the mean log_{10} body size (mm) while the shaded area represents the range of predictions. 17

18





- Table 1. Summary of the generalised linear model describing the probability of validity for Ichneumonids in our dataset. The model had binomial errors and a logit link. Bootstrapped estimates (1000 iterations, using 90% of the data) of the coefficients are given along with their 95% confidence intervals. Current validity was coded as 1 for valid and 0 for not valid. The alpha value was set at 0.001, thus every term retained in the model has a p-value of
- 6 <0.001. Cohen's Kappa for the model was 0.523 (95% CI= 0.001) and the area under the
- 7 ROC curve was 0.886 (95% CI=<0.001).

	Coefficient	
Main Effect Terms	Estimate	95% CI
(Intercept)	20.541	0.033
\log_{10} (body length in mm)	-0.901	0.010
Time known (yrs)	0.015	< 0.001
Continent (Asia-Temperate)	2.575	0.049
Continent (Asia-Tropical)	2.210	0.053
Continent (Australasia)	-11.946	0.058
Continent (Europe)	-5.906	0.036
Continent (Northern America)	15.049	0.053
Continent (Pacific)	879.272	1118.622
Continent (Southern America)	-5.784	0.054
log ₁₀ (area in km ²)	-2.944	0.005
Interaction terms	0.4.00	0.040
log ₁₀ (body length) : Continent (Asia-Temperate)	0.102	0.013
log ₁₀ (body length) : Continent (Asia-Tropical)	-1.719	0.018
\log_{10} (body length) : Continent (Australasia)	-1.273	0.023
log ₁₀ (body length) : Continent (Europe)	0.654	0.011
\log_{10} (body length) : Continent (Northern America)	-0.024	0.012
log ₁₀ (body length) : Continent (Pacific)	-88.740	211.787
log ₁₀ (body length) : Continent (Southern America)	-0.512	0.016
Time known : Continent (Asia-Temperate)	-0.011	< 0.001
Time known : Continent (Asia-Tropical)	-0.014	< 0.001
Time known : Continent (Australasia)	-0.017	< 0.001
Time known : Continent (Europe)	-0.012	< 0.001
Time known : Continent (Northern America)	-0.006	< 0.001
Time known : Continent (Pacific)	0.011	0.045
Time known : Continent (Southern America)	-0.018	< 0.001
Continent (Asia-Temperate) : log ₁₀ (Area)	-0.022	0.007
Continent (Asia-Tropical) : log ₁₀ (Area)	0.273	0.008
Continent (Australasia) : log ₁₀ (Area)	2.189	0.008
Continent (Europe) : log10 (Area)	1.026	0.006
Continent (Northern America) : log10 (Area)	-1.601	0.008
Continent (Pacific) : log10 (Area)	-169.496	188.414
Continent (Southern America) : log10 (Area)	1.326	0.008

Jones, O.R., Purvis, A, and Quicke, D.J.L., Latitudinal gradients in overdescription rate affect
 macroecological inferences using species list data. Ecography.

Supplementary material: Appendix 1





Figure A1. Map depicting the continental scale areas and regions used in this study. Each region is made up of a number of subregions, detailed in Table A1. From Brummitt 2001.





4

Figure A2. Macroecological patterns in the Ichneumonidae in different longitudinal windows (-165°--30°, -15°-60° and 60°-180°). (A-C) Species richness patterns. The solid black line indicates current richness. The broken line shows predicted richness after accounting for

overdescription. Prediction range is indicated by the shaded area. (D-F) Proportion taxonomic
 overdescription. The line represents mean predicted synonymy, after correcting for

- overdescription. The line represents mean predicted synonymy, after correcting for
 overdescription. Shaded area indicates the range of the predictions. (H-J) Body size. The line
- represents mean \log_{10} body size (mm) and the shaded area shows the range of predictions.





1 2 Figure A3. Macroecological patterns in the Ichneumonidae within different latitudinal bin 3 sizes (5° and 20°, left panel and right panel respectively). (A-B) Global species richness 4 patterns. The solid black line indicates current species richness. The broken line indicates 5 predicted richness after accounting for taxonomic overdescription. The shaded area indicates 6 the range of the predictions. (C-D) Proportion taxonomic overdescription. The line represents the mean predicted synonymy, after correcting for overdescription, while the shaded area 7 8 indicates the range of the predictions. (E-F) Species body size. The line represents the mean log_{10} body size (mm) while the shaded area represents the range of predictions. 9 10

- 1 Table A1. 'Continents' (in bold type), regions and, in parenthesis, subregions that were used in this
- 2 study. Definitions used come from Brummitt 2001.

3 Africa

- 4 East Tropical Africa (Kenya, Tanzania, Uganda)
- 5 Macaronesia (Azores, Canary Is., Madeira)
- 6 Middle Atlantic Ocean (Ascension, St.Helena)
- 7 Northeast Tropical Africa (Chad, Eritrea, Ethiopia, Socotra, Somalia, Sudan)
- 8 Northern Africa (Algeria, Egypt, Libya, Morocco, Tunisia, Western Sahara)
- 9 South Tropical Africa (Angola, Malawi, Mozambique, Zambia, Zimbabwe)
- 10 Southern Africa (Botswana, Lesotho, Namibia, Swaziland)
- 11 West Tropical Africa (Benin, Burkina, Gambia, Ghana, Guinea, Ivory Coast, Liberia, Mali,
- 12 Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Togo)
- 13 West-Central Tropical Africa (Burundi, Cameroon, Central African Republic, Congo,
- 14 Equatorial Guinea, Gabon, Rwanda)
- 15 Western Indian Ocean (Comoros, Madagascar, Mauritius, Reunion, Rodrigues, Seychelles)

16 Antarctic

17 Subantarctic Islands (Falkland Is.)

18 Asia-Temperate

- 19 Arabian Peninsula (Gulf States, Kuwait, Oman, Saudi Arabia, Yemen)
- 20 Caucasus (North Caucasus, Transcaucasus)
- 21 China (China North-Central, China South-Central, China Southeast, Hainan, Inner
- 22 Mongolia, Manchuria, Qinghai, Tibet, Xinjiang)
- 23 Eastern Asia (Japan, Korea, Nansei-shoto, Ogasawara-shoto, Taiwan)
- 24 Middle Asia (Kazakhstan, Tadzhikistan, Turkmenistan, Uzbekistan)
- 25 Mongolia (Mongolia)
- 26 Russian Far East (Amur, Kamchatka, Khabarovsk, Magadan, Primorye, Sakhalin)
- 27 Siberia (Altay, Buryatiya, Chita, Irkutsk, Krasnoyarsk, Tuva, West Siberia, Yakutskiya)
- 28 Western Asia (Afghanistan, Cyprus, Iran, Iraq, Lebanon-Syria, Palestine, Turkey)

29 Asia-Tropical

- 30 Indian Subcontinent (Bangladesh, Chagos Archipelago, East Himalaya, India, Maldives,
- 31 Nepal, Pakistan, Sri Lanka)
- 32 Indo-China (Cambodia, Laos, Myanmar, Nicobar Is., Thailand, Vietnam)
- 33 Malesia (Borneo, Jawa, Lesser Sunda Is., Malaya, Maluku, New Guinea, Philippines,
- 34 Sulawesi, Sumatera)

35 Australasia

- Australia (New South Wales, Norfolk Is., Northern Territory, Queensland, South Australia,
 Tasmania, Victoria, Western Australia)
- 38 New Zealand (New Zealand South)

39 Europe

- 40 Eastern Europe (Baltic States, Belarus, Central European Russia, East European Russia,
- 41 North European Russia, Northwest European Russia, South European Russia, Ukraine)
- 42 Middle Europe (Austria, Belgium, Czechoslovakia, Germany, Hungary, Netherlands,
- 43 Poland, Switzerland)
- 44 Northern Europe (Denmark, Faeroe Islands, Finland, Great Britain, Iceland, Ireland,
- 45 Norway, Svalbard, Sweden)
- 46 Southeastern Europe (Albania, Bulgaria, Greece, Italy, Kriti, Romania, Sicilia, Yugoslavia)
- 47 Southwestern Europe (Baleares, Corse, France, Portugal, Sardinia, Spain)

1 Northern America

- 2 Eastern Canada (New Brunswick, Newfoundland, Nova Scotia, Ontario, Prince Edward I.,
- 3 Quebec)
- 4 Mexico (No subregions in dataset)
- 5 North-Central U.S.A. (Illinois, Iowa, Kansas, Minnesota, Missouri, Nebraska, North
- 6 Dakota, Oklahoma, South Dakota, Wisconsin)
- 7 Northeastern U.S.A. (Connecticut, Indiana, Maine, Massachusetts, Michigan, New
- 8 Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode I., West Virginia)
- 9 Northwestern U.S.A. (Colorado, Idaho, Montana, Oregon, Washington, Wyoming)
- 10 South-Central U.S.A. (New Mexico, Texas)
- 11 Southeastern U.S.A. (Alabama, Arkansas, Delaware, District of Columbia, Florida,
- 12 Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina,
- 13 Tennessee, Virginia)
- 14 Southwestern U.S.A. (Arizona, California, Nevada, Utah)
- 15 Subarctic America (Alaska, Greenland, Northwest Territories, Nunavut, Yukon)
- 16 Western Canada (Alberta, British Columbia, Manitoba, Saskatchewan)

17 Pacific

- 18 North-Central Pacific (Hawaii)
- 19 Northwestern Pacific (Caroline Is., Marianas, Marshall Is.)
- 20 South-Central Pacific (Caroline Is., Christmas I., Cook Is., Easter Is., Society Is.)
- 21 Southwestern Pacific (Fiji, New Caledonia, Samoa, Solomon Is., Tonga, Vanuatu)

22 Southern America

- 23 Caribbean (Bahamas, Bermuda, Cuba, Dominican Republic, Haiti, Jamaica, Leeward Is.,
- 24 Netherlands Antilles, Puerto Rico, Trinidad-Tobago, Windward Is.)
- 25 Central America (Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua,
- 26 Panama)
- 27 Northern South America (French Guiana, Guyana, Suriname, Venezuela)
- 28 Southern South America (Argentina Northeast, Chile North, Juan Fernandez Is., Paraguay,
- 29 Uruguay)
- 30 Western South America (Bolivia, Colombia, Ecuador, Galapagos, Peru)
- 31

 Table A2. Current and predicted species richness values across different latititudes, globally, from Figure 2A.

		Predictions from Monte Carlo simulation			
Latitude	Current species richness	Minimum	Average	Maximum	
-85° to -75°	0	0	0	0	
-75° to -65°	0	0	0	0	
-65° to -55°	0	0	0	0	
-55° to -45°	0	0	0	0	
-45° to -35°	170	71	82.1	90	
-35° to -25°	696	263	281.9	304	
-25° to -15°	1109	618	645.3	673	
-15° to -5°	1952	1139	1173.7	1210	
-5° to 5°	2444	1597	1643.3	1676	
5° to 15°	3540	1720	1764.8	1797	
15° to 25°	1650	1125	1152.3	1179	
25° to 35°	4140	2656	2708.7	2761	
35° to 45°	9313	5832	5917.8	5989	
45° to 55°	8816	5655	5731.9	5812	
55° to 65°	7375	4536	4618.2	4711	
65° to 75°	5667	3512	3602.8	3683	
75° to 85°	133	81	92.8	105	

1 Table A3. Species richness and ranked species richness for Ichneumonidae at the regional

2 level. The table is sorted by current rank. Predicted mean and variance come from 1000

3 Monte Carlo simulations. Region definitions used come from Brummitt 2001.

	Curront	Predicted Predicted		Current	Prodicted	
Region	Divorsity	Diversity	Diversity	Donk	I reulcieu Donk	
	Diversity	(Mean)	Variance	Nalik	Nalik	
Middle Europe	4878	3451	351.2	50	50	
Eastern Europe	3552	2049	477.4	49	47	
Northern Europe	3439	2285	308.1	48	49	
Southwestern Europe	3096	2067	292.1	47	48	
Southeastern Europe	2963	2043	269.8	46	46	
Northeastern U.S.A.	2323	1384	171.3	45	44	
Eastern Asia	2198	1423	185.7	44	45	
Eastern Canada	2063	1183	185.1	43	42	
Western Canada	1928	1107	188.5	42	40	
China	1816	1061	188.4	41	39	
Northwestern U.S.A.	1777	1013	157.0	40	38	
Malesia	1772	1129	71.6	39	41	
Southeastern U.S.A.	1650	1008	116.3	38	37	
Southwestern U.S.A.	1573	889	112.5	37	36	
Indian Subcontinent	1571	1212	121.6	36	43	
Central America	1374	335	17.9	35	20	
Russian Far East	1362	653	203.5	34	34	
North-Central U.S.A.	1348	795	108.9	33	35	
Siberia	944	460	152.1	32	26	
Subarctic America	934	532	94.9	31	31	
Mexico	888	464	24.5	30	27	
Western Asia	875	558	79.1	29	33	
Caucasus	873	483	118.4	28	28	
South-Central U.S.A.	813	486	54.2	27	29	
Brazil	744	391	77.7	26	22	
Indo-China	692	556	48.2	25	32	
East Tropical Africa	678	518	48.4	24	30	
Middle Asia	636	363	70.3	23	21	
Western South America	624	392	27.8	22	23	
Western Indian Ocean	595	450	21.0	21	25	
Southern South America	561	264	25.8	20	17	
Southern Africa	553	420	40.3	19	24	
Northern Africa	486	319	46.4	18	18	
W-Central Trop. Africa	454	324	51.2	17	19	
Australia	424	133	34.7	16	11	
Mongolia	412	213	53.8	15	16	
Northern South America	332	171	14.3	14	13	
West Tropical Africa	277	194	19.8	13	15	
South Tropical Africa	261	188	23.4	12	14	
Macaronesia	215	156	14.3	11	12	
Caribbean	203	108	6.6	10	10	
Southwestern Pacific	162	90	5.4	9	9	
Northeast Tropical Africa	128	83	12.5	8	8	
New Zealand	87	52	5.1	7	7	

North-Central Pacific	68	50	3.0	6	6
Northwestern Pacific	36	13	1.5	5	4
Arabian Peninsula	33	22	2.6	4	5
South-Central Pacific	20	13	0.5	3	3
Middle Atlantic Ocean	5	4	0.4	2	2
Subantarctic Islands	0	0	0.0	1	1

1 Table A4. Species richness and ranked species ricness for Ichneumonidae at the subregion level.

The table is sorted by current rank. Predicted mean and variance come from 1000 Monte Carlo
 simulations. Subregion definitions used come from Brummitt 2001.

		Predicted	Predicted		
	Current	Diversity	Diversity	Current	Predicted
Region	Diversity	Mean	Variance	Rank	Rank
Germany	3865	2628.8	308.9	284	284
France	2737	1777.0	271.7	283	283
Poland	2450	1658.1	263.9	282	282
Sweden	2402	1602.5	263.4	281	281
Finland	2278	1505.9	249.6	280	280
Austria	2216	1498.7	224.5	279	279
Great Britain	2169	1457.6	234.8	278	278
Hungary	1929	1304.9	207.5	277	277
Czechoslovakia	1762	1150.5	209.6	276	275
Romania	1726	1169.3	176.2	275	276
Bulgaria	1511	989.4	176.6	274	273
Quebec	1480	800.1	141.0	273	265
Japan	1475	933.0	114.1	272	272
Ontario	1455	843.4	137.5	271	270
India	1438	1100.8	116.5	270	274
New York	1408	774.4	108.8	269	264
Italy	1381	865.6	132.4	268	271
Baltic States	1306	821.0	158.0	267	268
British Columbia	1290	723.7	135.7	266	261
Netherlands	1271	805.4	145.7	264	267
Spain	1271	824.3	131.6	264	269
Belgium	1255	802.5	151.9	263	266
Ukraine	1195	730.2	147.2	262	262
Norway	1182	743.8	150.0	261	263
Michigan	1178	631.9	95.5	260	258
Costa Rica	1160	171.6	12.8	259	186
Alberta	1151	653.5	122.7	258	259
California	1126	615.4	88.5	257	256
Switzerland	1110	696.9	119.4	256	260
Colorado	1108	616.0	102.4	255	257
Maine	916	522.4	95.0	254	255
Northwest European					
Russia	851	468.6	139.0	253	252
Oregon	820	427.0	78.1	252	245
Central European Russia	817	442.6	129.6	251	247
Massachusetts	796	447.3	67.5	250	248
New Jersey	793	462.2	58.3	249	251
Pennsylvania	782	450.9	67.1	248	249
North Carolina	767	440.1	65.8	247	246
New Hampshire	765	412.8	72.5	246	243
Taiwan	746	499.1	80.0	245	254
Washington	744	402.5	76.0	244	241
Georgia	709	452.9	53.9	243	250
Maryland	704	406.3	55.1	242	242
Ohio	699	397.7	60.3	241	240

Alaska	682	371.2	77.1	240	235
Transcaucasus	676	377.3	84.6	239	238
Primorye	668	306.1	99.0	238	228
Arizona	651	376.7	48.6	237	237
Virginia	641	374.3	51.4	236	236
China Southeast	627	360.5	66.9	235	234
Turkey	610	379.4	61.7	234	239
Myanmar	589	489.0	40.9	233	253
Sakhalin	585	276.8	90.6	232	218
Idaho	570	310.7	57.4	231	230
Madagascar	558	424.9	20.5	230	244
Saskatchewan	555	298.2	56.7	229	225
Minnesota	553	294.2	55.5	228	222
Philippines	551	326.9	19.9	227	232
Belarus	546	296.1	73.0	226	223
Connecticut	533	301.5	50.0	225	227
Denmark	512	264.3	51.3	224	215
Manitoba	511	274.5	55.6	223	217
South Carolina	510	301.4	43.3	222	226
Texas	510	316.7	32.4	222	231
Illinois	498	308.1	44.7	220	229
Yugoslavia	490	296.8	52.1	219	224
Korea	477	274.3	44.8	218	216
New Guinea	454	193.3	12.0	217	196
Corse	442	277.8	46.4	216	220
Nova Scotia	434	229.1	44.9	215	209
Ireland	433	255.0	57.4	214	213
North European Russia	432	211.7	60.0	212	203
Tanzania	432	336.0	393	212	233
Kansas	423	255.9	33.6	211	214
China South-Central	422	195.6	43.1	210	199
Wisconsin	419	225.5	37.9	209	208
Newfoundland	418	236.7	49.6	208	211
Mongolia	412	213.1	53.8	200	204
New Mexico	409	232.7	34.4	206	210
Florida	404	232.7	25.3	200	210
Iawa	398	287.6	25.5	203	221
New Brunswick	392	216.1	40.3	201	205
West Virginia	392	210.1	38.9	202	203
South European Russia	386	186.4	59.3	202	193
North Caucasus	385	190.4	59.6	201	195
Congo	369	252.8	18.3	199	212
Vukon	362	176 4	40.5	108	102
I uKoli Kazakhstan	361	102.0	44.4 51.0	198	192
Nazakiistaii	250	200.9	31.0 28.4	197	201
Niomana Rhada I	241	200.8	38.4 22.1	190	201
	241 241	1/4./	33.1 20.6	194	190
I CHIICSSEC	241 227	221.U 106.6	30.0 17 2	194	200
r ciu Litab	<i>321</i> 210	190.0	1/.2	195	200
Utan Waxamin a	319	1/3.1	30.4 24 2	192	191
wyoming	51/	104.8	54.2	191	185

Khabarovsk	315	141.0	52.4	190	176
Argentina Northeast	312	148.7	17.4	189	179
Kamchatka	309	156.0	50.2	188	180
Manchuria	305	173.7	43.2	187	188
Missouri	303	188.4	27.1	186	194
Chita	302	146.7	46.9	185	178
Borneo	301	195.1	14.4	184	198
West Siberia	300	136.2	41.0	183	171
Northwest Territories	298	156.2	37.6	182	181
District of Columbia	283	174.5	23.6	181	189
Irkutsk	282	138.3	44.8	180	174
Sulawesi	265	203.0	13.8	179	202
Tadzhikistan	260	141.8	25.8	178	177
East European Russia	258	137.2	43.0	177	172
Greece	257	172.6	26.0	176	187
Yakutskiya	254	113.6	39.9	175	160
China North-Central	253	131.2	32.6	174	167
South Dakota	247	138.1	25.1	173	173
Queensland	246	68.0	16.5	172	132
Chile North	238	89.0	5.7	170	146
Uganda	238	160.7	20.4	170	182
Iowa	233	130.5	21.8	169	165
Louisiana	226	161.4	18.3	168	184
Panama	226	128.4	6.3	168	163
Nevada	225	135.7	19.5	166	170
Kenya	224	161.3	19.3	165	183
Nansei-shoto	222	104.1	12.5	164	154
Palestine	222	140.7	19.8	164	175
Algeria	213	129.0	25.0	162	164
Kentucky	200	123.9	17.6	161	162
Sri Lanka	194	133.5	14.9	160	168
Sumatera	191	131.1	14.9	159	166
Arkansas	185	135.3	16.4	158	169
Alabama	184	110.6	15.4	157	159
Indiana	181	103.6	17.4	156	153
Tunisia	179	118.5	16.0	155	161
Sicilia	177	106.5	18.7	154	155
Prince Edward I	169	81.8	14.9	153	142
Ecuador	165	101.3	7.3	152	150
Bolivia	161	107.1	8.7	150	156
Morocco	161	101.3	16.6	150	151
Nebraska	161	96.3	14.7	150	148
Guatemala	160	110.2	6.0	148	158
Krasnovarsk	157	76.1	26.0	147	136
Venezuela	155	81.1	99	146	140
Nepal	154	109.1	13.2	145	157
New South Wales	153	45.3	10.1	144	112
Guvana	146	76.9	61	143	137
Canary Is	143	101.4	11.0	142	152
Burvativa	142	63.9	22.5	141	126
	÷ ·-				

0.1.1:	1.4.1	0(1	0.0	1.40	1.4.7
Colombia	141	86.1	8.2	140	145
Uzbekistan	141	84.9	14.5	140	143
North Dakota	13/	/2.4	15.1	138	134
Angola	136	97.7	13.8	13/	149
Delaware	131	/5.5	12.4	136	135
Egypt	130	//.0	13.7	135	138
Turkmenistan	129	81.8	12.5	134	141
	126	65.8	4.4	133	129
Iran	118	/8./	11.0	132	139
Mississippi	118	85.1	8.6	132	144
Maluku	117	65.6	6.2	130	128
Sardinia	115	71.5	11.1	129	133
Guinea	112	89.9	6.8	128	147
Victoria	105	37.9	8.9	127	104
Altay	103	46.5	15.6	126	115
Ethiopia	102	67.1	10.0	125	131
Portugal	97	51.6	8.6	124	119
Amur	96	44.7	16.0	122	111
Thailand	96	56.7	7.4	122	123
Greenland	93	67.0	7.8	120	130
Tasmania	93	35.4	7.8	120	100
Zimbabwe	93	62.6	9.0	120	125
Afghanistan	92	54.3	9.6	118	121
Iceland	89	51.6	8.5	117	118
Paraguay	88	59.3	6.5	116	124
Madeira	87	64.8	5.8	114	127
New Zealand South	87	52.0	5.1	114	120
Pakistan	85	56.7	7.5	113	122
Western Australia	84	25.4	7.6	112	87
Tibet	82	39.4	8.7	111	106
Xinjiang	81	42.5	9.9	110	108
Solomon Is.	79	36.0	3.2	109	102
Magadan	72	29.4	11.2	108	93
Malava	72	43.4	5.2	108	109
Nunavut	71	45.7	7.8	106	113
Oklahoma	70	46.0	6.4	105	114
Albania	69	35.5	73	104	101
Cameroon	68	47.1	6.1	102	116
Hawaii	68	50.1	3.0	102	117
Suriname	68	36.5	3.2	102	103
Nigeria	65	31.3	7.0	100	97
Hainan	61	28.4	53	99	91
Oinghai	58	20.1 41.1	73	98	107
Togo	58	44.2	4.2	98	110
Inner Mongolia	56	27.2	7.2	96	88
Vietnam	55	32 4	5.7	90	08
Honduras	53	32. 4 31.2	3.7 2.4	95	70 06
Mozambique	55	21.2 21.1	∠. 4 6 1	7 4 02	90 00
New Caledonia	50	34.4	0.1	93 02	99 05
Divende	50	20.4 20 7	1.0	92	93 105
rwanua	4/	38./	4.2	91	105

0 1	1.0	20.5	1.0	0.0	0.4
Senegal	46	29.5	4.0	90	94
Baleares	45	27.9	2.6	88	89
Cyprus	45	24.4	4.1	88	85
Faeroe Islands	42	28.3	5.2	87	90
Nicaragua	41	22.3	2.7	86	82
Vanuatu	41	23.8	1.7	86	83
Central African Republic	40	19.5	5.9	84	79
Ivory Coast	39	18.6	5.0	82	77
Lebanon-Syria	39	29.0	3.1	82	92
Northern Territory	39	8.5	1.8	82	48
French Guiana	37	12.9	0.8	80	61
South Australia	35	13.9	3.6	79	66
Uruguay	34	19.3	2.4	78	78
Windward Is.	34	24.7	2.0	78	86
Ghana	32	14.1	2.9	75	67
Puerto Rico	32	17.4	2.4	75	75
Sierra Leone	32	20.1	2.7	75	80
Namibia	31	23.9	2.2	73	84
Fiji	30	16.3	2.1	72	73
Jamaica	29	16.1	2.0	70	71
Laos	29	15.9	31	70	70
Trinidad-Tobago	28	96	2.2	69	52
Belize	27	11.5	2.3	68	56
Caroline Is	27	9.5	0.9	68	51
Bangladesh	26	13.4	2.4	66	63
Sudan	26	12.0	3.5	66	59
Burundi	20	21.1	2.0	64	81
Gabon	25	16.8	2.0	64	74
Dominican Republic	23	0.1	1.0	62	/4
Libya	24	9.1 13-3	1.9	62	49 62
Lioya Lassar Sunda Is	24	13.3	5.5 1 7	50	65
Malawi	23	13.7	1.7	59	05
Nialawi Saudi Anabia	23	18.0	2.0	59	/6
	23	13.0	1.0	59	09
Azores	20	10.2	1.0	50	12
Mauritius	20	10.2	1.4	56	53
Reunion	19	14.2	1.1	54	68
Seychelles	19	12.3	1.0	54	60
Svalbard	19	13.4	1.9	54	64
Tuva	18	7.2	3.1	52	44
Samoa	16	11.8	0.8	50	58
Zambia	16	7.0	1.7	50	43
El Salvador	15	7.6	1.2	48	46
Society Is.	15	11.3	0.4	48	55
East Himalaya	14	8.1	0.9	45	47
Equatorial Guinea	14	11.8	0.9	45	57
Eritrea	14	10.9	0.9	45	54
Liberia	14	9.4	1.6	45	50
Marianas	14	4.9	1.0	45	39
Bermuda	12	6.1	0.6	41	41
Kriti	12	7.3	1.5	41	45

Yemen	12	6.7	1.1	41	42
Haiti	9	3.6	0.3	38	31
Leeward Is.	9	3.9	0.7	38	34
Bahamas	8	4.5	0.9	37	38
Juan Fernandez Is.	7	5.6	0.6	36	40
Niger	7	4.2	0.6	36	35
Galapogos	6	0.0	0.0	33	3
Iraq	6	4.3	0.5	33	36
Ogasawara-shoto	6	0.7	0.2	33	8
Benin	5	1.7	0.2	30	21
Botswana	5	2.6	0.3	30	27
St.Helena	5	4.4	0.4	30	37
Burkina	4	1.9	0.5	26	24
Lesotho	4	3.6	0.4	26	32
Nicobar Is.	4	3.8	0.2	26	33
Norfolk Is.	4	0.6	0.2	26	7
Somalia	4	2.3	0.5	26	26
Chad	3	1.8	0.7	20	22
Chagos Archipelago	3	1.2	0.5	20	17
Comoros	3	2.8	0.2	20	29
Gambia	3	2.7	0.2	20	28
Marshall Is.	3	1.1	0.5	20	16
Oman	3	2.8	0.2	20	30
Ascension	2	1.5	0.3	14	18
Cambodia	2	0.0	0.0	14	3
Easter Is.	2	1.5	0.3	14	18
Rodrigues	2	1.6	0.2	14	20
Socotra	2	2.0	0.0	14	25
Tonga	2	1.9	0.1	14	23
Christmas I.	1	0.9	0.1	6	13
Cook Is.	1	0.9	0.1	6	12
Gulf States	1	0.8	0.1	6	11
Kuwait	1	0.5	0.2	6	6
Maldives	1	0.0	0.0	6	3
Mali	1	0.0	0.0	6	3
Mauritania	1	0.9	0.1	6	14
Netherlands Antilles	1	0.7	0.2	6	9
Swaziland	1	0.8	0.2	6	10
Western Sahara	1	1.0	0.0	6	15
Falkland Is.	0	0.0	0.0	1	3