

RESEARCH ARTICLE

Predicting the distribution of poorly-documented species, Northern black widow (*Latrodectus variolus*) and Black purse-web spider (*Sphodros niger*), using museum specimens and citizen science data

Yifu Wang^{1*}, Nicolas Casajus², Christopher Buddle¹, Dominique Berteaux², Maxim Larrivée³

1 Department of Natural Resource Sciences, McGill University, Montreal, Quebec, Canada, **2** Canada Research Chair on Northern Biodiversity, Université du Québec à Rimouski, Rimouski, Québec, Canada, **3** Insectarium de Montréal, Montréal, Québec, Canada

* yifu.wang@mail.mcgill.ca



OPEN ACCESS

Citation: Wang Y, Casajus N, Buddle C, Berteaux D, Larrivée M (2018) Predicting the distribution of poorly-documented species, Northern black widow (*Latrodectus variolus*) and Black purse-web spider (*Sphodros niger*), using museum specimens and citizen science data. PLoS ONE 13(8): e0201094. <https://doi.org/10.1371/journal.pone.0201094>

Editor: Matjaž Kuntner, National Institute of Biology, SLOVENIA

Received: January 26, 2018

Accepted: July 9, 2018

Published: August 8, 2018

Copyright: © 2018 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Predicting species distributions requires substantial numbers of georeferenced occurrences and access to remotely sensed climate and land cover data. Reliable estimates of the distribution of most species are unavailable, either because digitized georeferenced distributional data are rare or not digitized. The emergence of online biodiversity information databases and citizen science platforms dramatically improves the amount of information available to establish current and historical distribution of lesser-documented species. We demonstrate how the combination of museum and online citizen science databases can be used to build reliable distribution maps for poorly documented species. To do so, we investigated the distribution and the potential range expansions of two north-eastern North American spider species (Arachnida: Araneae), the Northern black widow (*Latrodectus variolus*) and the Black purse-web spider (*Sphodros niger*). Our results provide the first predictions of distribution for these two species. We also found that the Northern black widow has expanded north of its previously known range providing valuable information for public health education. For the Black purse-web spider, we identify potential habitats outside of its currently known range, thus providing a better understanding of the ecology of this poorly-documented species. We demonstrate that increasingly available online biodiversity databases are rapidly expanding biogeography research for conservation, ecology, and in specific cases, epidemiology, of lesser known taxa.

Introduction

Data deficiency is the main obstacle for developing accurate distribution maps [1]. For most known species, museum specimens, private collections and historical literature are the major

data sources to study species biogeography. Nevertheless, limited funding and the magnitude of the task to manually digitize museum specimen records result in few museums with digital, spatially-explicit databases, therefore limiting museum specimen data availability for the scientific community. Museums that have digital databases readily available sometimes lack information such as latitude, longitude, and precision of the GPS coordinates, which are necessary for generating species distribution models. Museum specimen data also often cover a limited time span [2, 3]. Observations or collections over the last 20 years are rarely found in museums, thus creating temporal gaps in database coverage. Such temporal gaps must be filled to produce reliable distribution predictions as recent anthropogenic climate and land use change hastens species range shifts [4, 5].

On the other hand, the internet allows convenient and fast data sharing, and this can help scientists rapidly gather research grade data with careful vetting. Currently, many citizen science (CS) initiatives not only provide natural history knowledge for the public, but also gather observation records of high quality [6]. These observations, in the form of collected specimens, photos, or sightings, are submitted to online databases by amateurs and researchers [7–13]. In platforms such as eBird and eButterfly, each observation submitted is vetted by regional experts who validate the precision of geographic coordinates, the observation date, and the species determination, sometimes correcting species identification when applicable. In other platforms such as iNaturalist and Bugguide, the observer community self-controls the quality of identifications (crowd-source identification) and possible misidentifications are questioned and moved to a special forum waiting for expert confirmation. Citizen Science platforms also often provide an open-access to all users. Many studies have been successful at using citizen science data in tracking species distribution, from birds to invasive plants [4–6, 14, 15]. Occurrence data originating from museum collections and citizen scientists are generally complementary in their geographic coverage, which often improves the predictive power of species distribution models [7–12, 16]. Combining museum with citizen science data could be an optimal choice to predict and determine ranges and range shifts over time for under-documented species.

Determining reliable species distributions is a complex process requiring large amounts of data and effort which used to be only possible for well documented taxa. Species range shifts due to global change also adds complexity to the determination of species distribution [17–19]. Temperature and precipitation generally drive the distribution of temperate arthropods, particularly when modelling temperate spider distributions [20–25]. More precisely, the minimum temperature of the coldest period directly relates to the capacity of surviving through the winter for the Brown recluse spider (*Loxosceles reclusa*) [26]. Warming winter temperatures and extension of the growing season into the fall may delay such constraints and allow longer active periods through the delay in the first killing frosts increasing the possibility of northward range expansion [27–29]. Summer temperature regimes relate to reproductive success of some spider species, for example *Latrodectus hasseltii* [23]. Moisture level represented by precipitation data is also influential for terrestrial arthropod distribution as inadequate or excessive precipitation may lead to excessive stress due to extreme variations in moisture regimes [23, 24, 30].

The mapping complex effect from climate change on species distribution is simplified by major developments in species distribution model (SDM) algorithms and user-friendly interfaces over the last decade which allowed researchers and conservation practitioners to build reliable distribution models [31–33]. The ability of some SDMs to handle presence-only data also enables the determination of more detailed ranges of less-known taxa as these models can work with limited and often spatially-biased occurrence records, and even with data carrying some spatial uncertainty [34, 35]. Such improvements in the performance of SDMs coupled

with the increasing availability of online research-grade distribution data greatly strengthen our capacity to predict the ranges of poorly documented taxa [16, 36].

Knowing the distribution of poorly documented taxa helps to understand their ecology as it provides information about where populations occur, along with their habitat requirements. This also potentially allows forecasting the species vulnerability to environmental change or human activities [37]. Determining potential habitat is also essential for planning species management strategies. Moreover, ecosystem management often uses the presence or absence of certain species as ecological indicators [38]. Poorly known species are sometimes assigned with inappropriate conservation status. New knowledge on their biogeography can prompt downgrading or upgrading of their conservation status [39]. In addition, some poorly known species are potential disease vectors or are venomous. Thus, knowing their range is crucial for people to apply corresponding management and public education plans.

We selected two poorly documented species of spiders: *Latrodectus variolus* (Walckenaer) (Northern black widow) and *Sphodros niger* (Hentz) (Black purse-web spider) to determine if reliable species distribution models can be made for poorly known species by combining museum and citizen science data. North black widow is a habitat generalist, it can be found from mesic to xeric deciduous forest and builds web high up in trees [40]. It also inhabits in human modified landscapes such as pine plantations and downed fence posts, building webs close to ground and in small burrows [41]. The black widow clade is known for the venomous bite of its species [42] and represents a human health concern throughout its range. Closely monitoring its distribution is thus important. The Black purse-web spider is cryptic, poorly known, and ranked as vulnerable in Virginia, US, yet its distribution is poorly understood as very few specimens have been collected [43]. *Sphodros niger* is also more habitat selective than many other spider species, preferring dry sandy/rocky woodland area [44–46]. Knowing more precisely its potential distribution would facilitate its management at regional levels by more reliably identifying its potential bioclimatic niches.

Another reason to choose these two species was due to the feasibility of species-level identification base on obvious features. Citizen science records are often collected by non-experts who do not know key features for species identification and only take photos of the whole animal. Obvious body features allow species-level identification to be feasible with such photos. More importantly, crowd-source identification may not be done by experts and misidentification is possible [47]. Thus, choosing species that can be accurately identified based on easily visible body features can simplify the process of re-validating crowd-source identification and insuring data accuracy.

The first objective of this research was to assess whether combining citizen science and museum data allows to successfully model the distribution of two poorly documented spider species. Our second objective was to determine the suitable bioclimatic niches of the two species and explore whether these species expanded north as recently documented for other taxa in North America [4, 48].

Materials and methods

Species data

We gathered distribution data from various sources including museums, research centers, literature, personal collections, and online citizen science projects [7, 8, 12]. Data collected from museums or institutes were accessed through open online database or through literature that cited these specimens (details see S1 and S2 Files), except records from Canadian National Collection which were collected through visiting this institute. To standardize the data collected online and to maximize validity of our dataset, we removed any records that could not be

Table 1. Summary of occurrence data available for *Sphodros niger* and *Latrodectus variolus* distribution models, including the sources of data, their period of collection, and sample sizes. Other sources include private collections, personal observations, and news articles.

Species	1960–1989 from museum, literature, and other sources	1990–2016 (1990–2015 for <i>S. niger</i>) from museum, literature, and other sources	1960–1989 from citizen sciences	1990–2016 (1990–2015 for <i>S. niger</i>) from citizen sciences	Total
<i>Sphodros niger</i>	44	39	0	14	97
<i>Latrodectus variolus</i>	22	47	0	95	164

<https://doi.org/10.1371/journal.pone.0201094.t001>

determined to species-level based on provided photos. *Latrodectus variolus* and *S. niger* can be both identified with a sufficiently high accurate rate by their unique body coloration and patterns. For *L. variolus*, the majority of both male and female have a disconnected hour-glass pattern on their ventral abdomen contrasting with the partially sympatric *L. mactans* to the south which mostly have complete hour-glass marks. Overall coloration of *L. variolus* also distinguish them from other *Latodectus* species and guide used for vetting can be found in McCrone and Levi [40], Jensen [49], and Wilson [41]. For *S. niger*, its big forward-projecting chelicerae distinguish it from spiders of other families. The overall black to dark reddish-brown coloration and stubby brown to black legs differentiate them from other Atypidae species (description see reference) [46, 50–52]. We also removed the ones that were suspected of being pet animals, for example photos taken in vivarium setting. When only locality information was available with no geographic coordinates, we calculated the geographic uncertainty of the given locations using Georeferencing Quick Reference Guide [53]. Records with uncertainty above 10 km were removed from analysis as their precision was beyond our grid cell size which is 10 km by 10 km. In total, 97 Black purse-web spider records and 164 Northern black widow records were used for modeling their distribution (Table 1).

Environmental data

Climate data were constructed with ANUSPLIN, a regression splines interpolation, using all available weather station data in North America [54, 55]. Climate data resolution is at 10 arc minute resolution annually from 1960 to 2010 and was divided into two time-periods: 1960–1989 and 1990–2010. Then climate data of the two periods were averaged respectively using raster calculator in ArcGIS to represent historical and current climate. These climate data included: annual mean temperature, minimum temperature of the coldest period, mean temperature of the warmest quarter, mean temperature of the coldest quarter, annual precipitation, and precipitation seasonality (coefficient of variance).

Species distribution models

Species distribution models were created using MaxEnt 3.3.3k [56, 57]. It is one of the best distribution model techniques using presence-only data [31, 34, 58]. MaxEnt is widely used to predict distribution of many taxon, including spider distribution [59]. MaxEnt requires two types of data for modeling distribution, appropriate environmental layers and species distribution data containing latitude and longitude.

We constructed the current distribution model using climate data between 1990 and 2010 and species occurrence data between 1990 and 2016 (1990 to 2015 for *S. niger*). It is common practice in species distribution modelling that observations obtained slightly outside of the environmental variable timeframe are included, especially when occurrence data are rare [16, 60]. The historical distribution model used both occurrence and climate data between 1960

and 1989. Background points are used in presence-only species distribution model like Max-Ent to model pseudo-absences for species. Background points ($n = 10,000$) were identified across the study area. Models were constructed using the “samples with data” approach, using 10-fold cross-validation of model outputs against held-back species observations and converted to binary predictions of presence/absence using a 0.39 threshold [61, 62]. The models were iterated 100 times and the mean of these model outputs was used as the consensus estimate of each species’ distribution across the modelling extent. The accuracy of our models was evaluated using the area under the receiver operating characteristic curve (AUC), correlation coefficients (COR), and true skill statistics (TSS) [34, 63–66]. Models with good predictive performance have an AUC value close to 1. COR show how well the predicted value fits real data [31]. High COR values indicate strong positive correlation between predictions and actual presences of the species, which proves good model performance with high confidence [31, 67]. TSS score between 0.40 and 0.75 shows good predictive performance of the model while TSS score above 0.75 shows excellent performance [64].

Since not all 10 km x 10 km grid cells were sampled for species presence in our study area, some spatial bias could affect model predictions [68–70]. Several methods can address this bias in species distribution models [35, 71, 72], and we followed the method used by Elith *et al.* [32], which up weights a grid cell with fewer neighbors in geographic space. We calculated the number of occurrences in a chosen neighborhood for each grid cell and weighted this number by a Gaussian function (see [71] for further details). The resulting bias grids showed higher weight in areas that were more intensively sampled. In the Gaussian function, the standard deviation parameter must be specific to the species. We followed the recommendations of Clements and Rayan [71] and defined this parameter as the diameter of a “circular” home range of the species. We then estimated the surface of the home range using the 95% kernel approach [73].

We tested the existence of potential range shifts in the two spider species by comparing the respective northern limit of modeled historical distribution and current distribution. A paired t-test was run to compare the difference between the latitudes of predicted northern limits during the two periods. We also calculated the mean latitudinal difference between the modeled northern limits to quantify the proposed range shifts. Average latitudinal differences and associated standard deviations were converted to distances (km) to facilitate result interpretation.

Results

We successfully modeled the current distributions of *Latrodectus variolus* (Fig 1) and *Sphodros niger* (Fig 2). The historical distribution of *S. niger* was modeled successfully (Fig 2) but we failed to predict the historical distribution for *L. variolus* which is not included in maps. All illustrated range predictions had very high AUC, COR and TSS values (Table 2).

The result of the *S. niger* 1990–2015 model indicates new bioclimatic niches north of the modeled 1960–89 distribution range (especially in Ontario, Canada), but also a contraction in the southwest of the modeled distribution range, especially in Arkansas, Missouri and Tennessee. Two recent occurrences originating from CS in Minnesota and Wisconsin suggest a possible broader current range to the northwest for *S. niger* (Fig 2). Our northern edge shift calculation was only performed in *S. niger* models as the historical range model of *L. variolus* was not successful. The northern limit comparison reveals that there is a significant difference between the two northern edges of the models ($n = 294$, $t = 5.751$, $df = 293$, $p\text{-value} < 0.001$). The predicted northern shift as the mean difference between the two edges was -35 km with a high SD of ± 107 km.

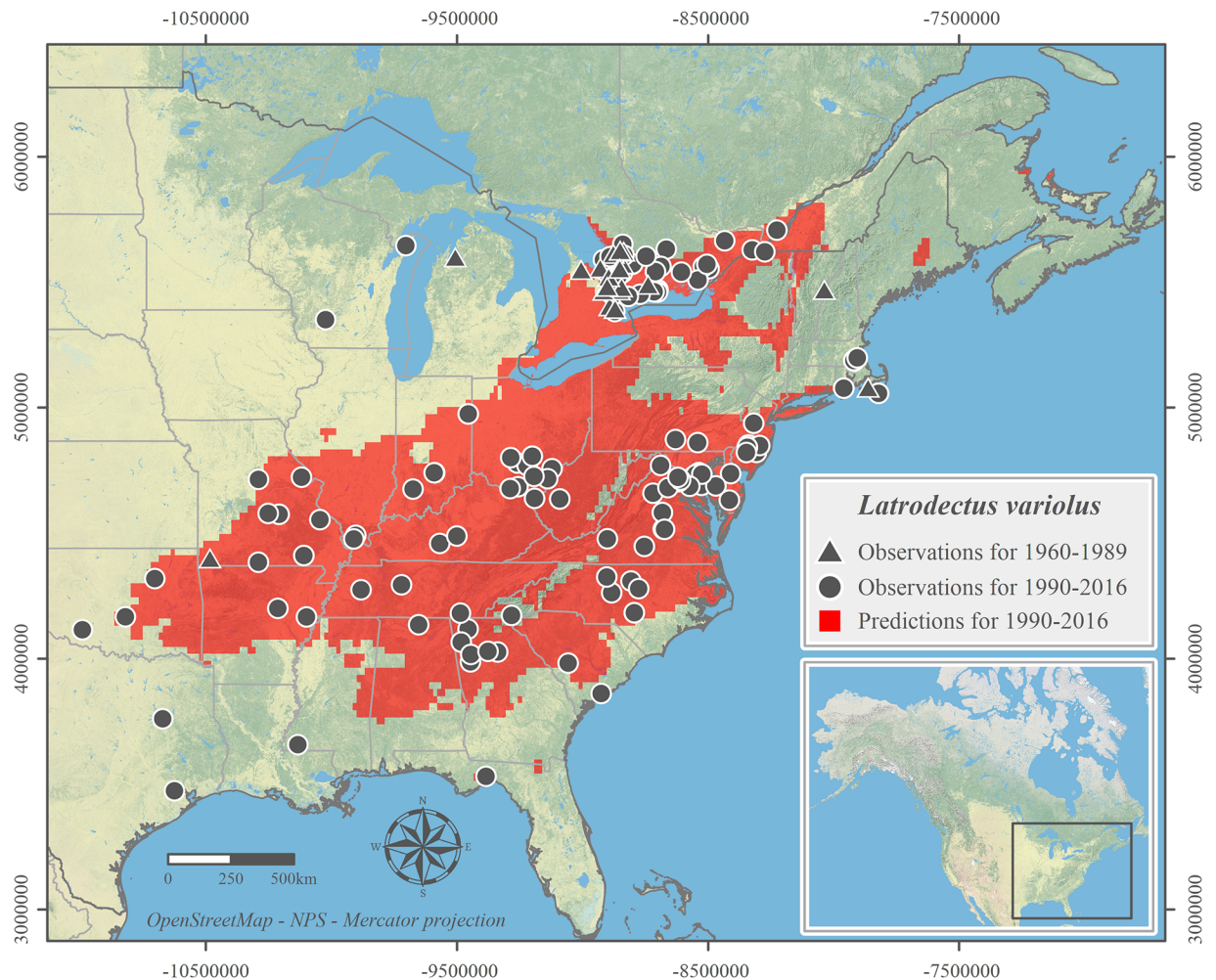


Fig 1. Suitable climatic habitat of *Latrodectus variolus* predicted from 1990–2016 observation records. Observation records are shown for both 1960–1989 and 1990–2016.

<https://doi.org/10.1371/journal.pone.0201094.g001>

The contribution of each environmental factor towards the final models was shown in the following table (Table 3). Mean temperature of the warmest quarter was the most important bioclimatic variable determining the current modeled range of *L. variolus* while mean temperature of the coldest quarter was the most important factor determining the modeled range of *S. niger* in both time periods (Table 3).

Discussion

We provide the first bioclimatic range predictions for two poorly known spider species, *Latrodectus variolus* and *Sphodros niger*. Our results show that constructing reliable species distribution models (very high AUC, COR and TSS values) is feasible for poorly documented taxa using multi-sourced data including well validated and curated online citizen science data. Phenological and range-shift responses for species in North America occurred mostly after 1960 [18], as did directional climate change, which is very likely due to human activities [74]. Therefore, 1960 was used as a historical baseline in which to construct our species distribution models. 67% of our *L. variolus* post 1990 records come from vetted online CS data (Table 1), which emphasizes the importance of this new data source in increasing our understanding of the

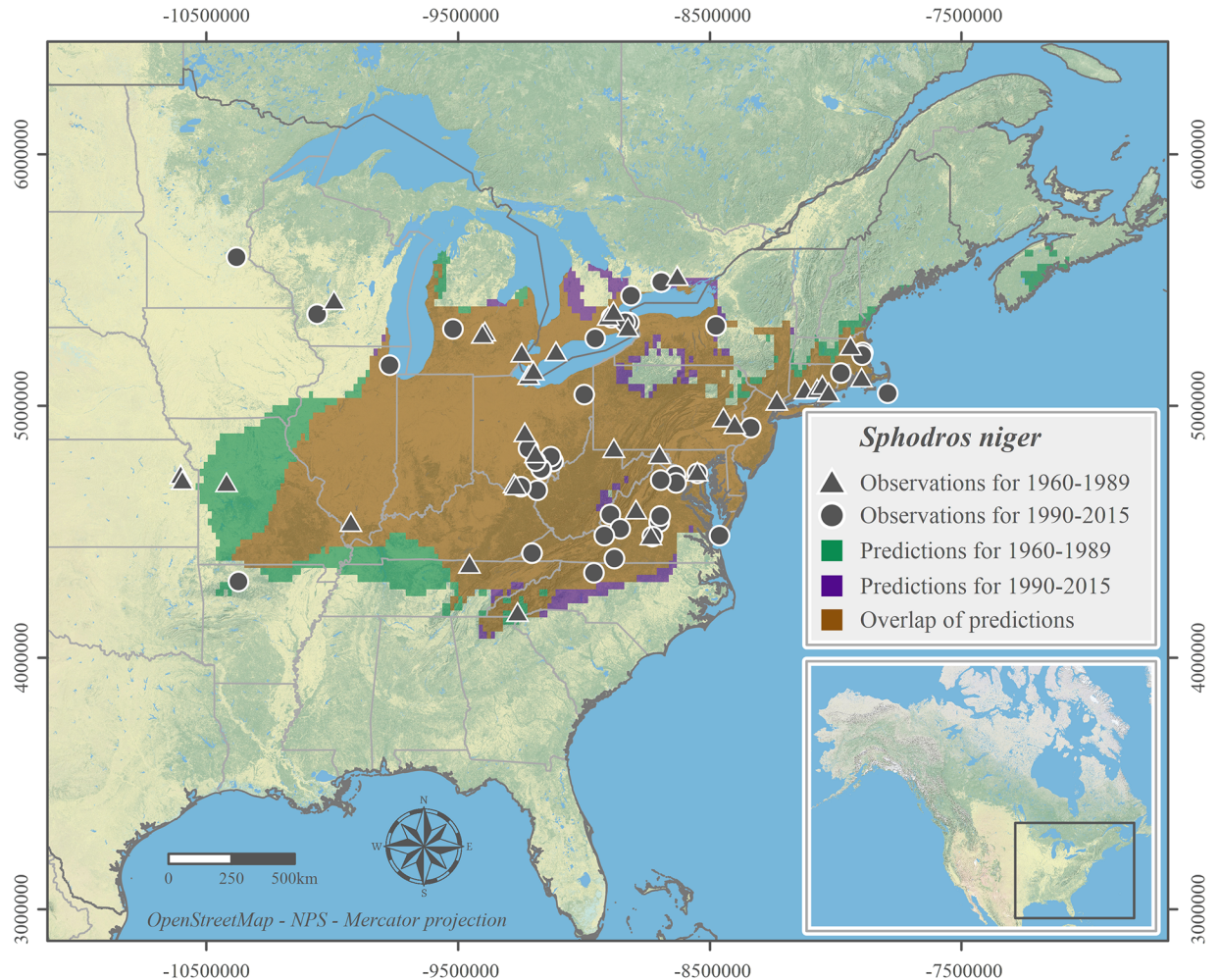


Fig 2. Suitable climatic habitat of *Sphodros niger* predicted from 1960–1989 and 1990–2015 observation records. Observation records are shown for both 1960–1989 and 1990–2015.

<https://doi.org/10.1371/journal.pone.0201094.g002>

distribution and biology of poorly known taxa. The failure to model the historical distribution range (1960–1990) of *L. variolus* is due to limited data records, 22 occurrences in total, which is too close to the minimum sample size for modeling wide-spread species [75]. This data deficiency induced failure further reinforces the necessity of digitalizing all available historical records and the incorporation of CS data for studying poorly-documented species. We successfully overcame the challenges described by Dickinson *et al.* [76], particularly the

Table 2. Predictive performance of *Latrodectus variolus* and *Sphodros niger* distribution models assessed through the receiver operating characteristic curve (AUC), correlation coefficients (COR), and true skill statistics (TSS). Values are the average (\pm SD) of 100 iterated models.

Model	AUC	COR	TSS
<i>Latrodectus variolus</i>			
1990–2016	0.948 \pm 0.009	0.384 \pm 0.021	0.709 \pm 0.086
<i>Sphodros niger</i>			
1960–1989	0.965 \pm 0.013	0.277 \pm 0.030	0.754 \pm 0.161
1990–2015	0.967 \pm 0.018	0.293 \pm 0.029	0.802 \pm 0.145

<https://doi.org/10.1371/journal.pone.0201094.t002>

Table 3. The importance of the six bioclimatic predictors used in habitat range models of *Latrodectus variolus* and *Sphodros niger*.

Model	Mean annual temperature (%)	Minimum temperature of the coldest month (%)	Mean temperature of the warmest quarter (%)	Mean temperature of the coldest quarter (%)	Total annual precipitation in ml (%)	Precipitation seasonality (%)
<i>Latrodectus variolus</i>						
1990–2010	1.4 ± 0.9	1.3 ± 1.2	67.0 ± 2.9*	2.4 ± 0.8	4.9 ± 1.4	23.0 ± 3.0
<i>Sphodros niger</i>						
1960–1989	14.7 ± 3.5	12.2 ± 10.6	23.1 ± 3.7	41.1 ± 15.4*	3.0 ± 1.5	5.8 ± 2.3
1990–2010	0.2 ± 1.0	29.2 ± 14.9	14.5 ± 4.5	39.0 ± 15.1*	9.6 ± 3.9	7.6 ± 3.5

The percentages with their associated standard deviations are the contribution of each factor to the final results.

* Asterisk mark highlights the most influential factor for each model.

<https://doi.org/10.1371/journal.pone.0201094.t003>

geographical biases inherent to CS data. We incorporated observer errors and spatial sampling bias through multilayered data vetting and constructing sampling bias grid [32, 71]. Our method provides a valuable template for similar studies in future.

The distribution of both species was strongly driven by temperature. Mean temperature of the warmest quarter is the driving factor of the distribution of *L. variolus* (Table 3). Summer temperature influences spiders’ life history in multiple aspects including breeding behavior, web construction, prey availability, growth rate, and dispersal behavior [77–80]. For *S. niger*, mean temperature of the coldest quarter is the most influential environmental factor in both models. Temperature is closely related to the winter survival rate of many spider species and thus often defines their northern distribution limits [26].

We were not able to provide statistically significant evidence to support the proposed north shift scenario as either the historical range model failed (*L. variolus*), or the range shift calculation had a high SD (*S. niger*) which devalued the result. Nevertheless, our model predictions of the current northern limits of both species are based on very conservative model thresholds and both predicted ranges extend beyond previously documented regions (Figs 1 and 2) [43]. For *L. variolus*, we found that the northern most observation for the 1990–2016 period (located in Quebec) was 94 km north of the northern most observation for 1960–1989 (located in Ontario). The predicted suitable climatic niche of *L. variolus* for the 1990–2016 extends another 50 km north to specimens having yet to be recorded at the northeast of Montréal, QC. The model of *L. variolus* is also in accordance with recent observations north of its previously known range. These observations in Eastern Ontario [81, 82] and Southern Québec (from Montreal Insectarium entomological enquiry services in 2012, 2015 and 2016) provide strong empirical support to our northern range expansion hypothesis. The Montreal Insectarium has been answering on average 1500 public entomological inquiries since 1990 and received all three inquiries pertaining to *L. variolus* in Quebec after 2012. Thus, these observations beyond the historically known northern limit were not the result of the multiplication of CS projects in recent years. Climate change, which influences seasonal temperature pattern, could be a strong contributor to the increase in occurrences of Northern black widow beyond their historical northern limit [17, 23, 83, 84]. Similar expansion pattern has been revealed in other arthropod species in the northeast of North America where our focal species ranges are located, for example the giant swallowtail butterfly (*Papilio cressphontes*) [85] and other conspicuous butterfly species [4, 48]. Another spider of health concern in North America, the Brown recluse (*Loxosceles reclusa*), was shown to potentially expand their distribution range northwards under future climate change scenarios [84]. Thus, these two spider species are also possible to

respond similarly to the climate change induced relaxation of limiting factors and have expanded northwards.

Quick and successful colonization to new habitat is particularly possible for *L. variolus* as this species is a habitat and prey generalist. Thus, it is more likely to survive in new environments after long-distance ballooning, the main dispersal method of spiders [78, 86–88]. *L. variolus* also has a higher metabolic rate compared to other theridiid species, allowing it to have larger clutch sizes and thus a higher reproduction rate [89]. Therefore, *L. variolus* has several of the key assets generally associated with good colonizers. For *S. niger* being a habitat specialist, colonizing new habitat beyond its currently known range extent might be slower and less efficient due to the random aspect of long distance ballooning.

Different requirement on habitat type may also be the explanation to the different most important environmental factor in predicting models (Table 3). *Sphodros niger* being a habitat specialist preferring dry sandy/rocky woodland area, escaping cold winter might be difficult while *L. variolus* can be found frequently in man-made shelters. As the result, mean temperature of the coldest quarter is found to be the most influential factor for mapping *S. niger* distribution. On the other hand, relatively higher metabolic rate of *L. variolus* comparing to other theridiid species might be the reason why mean temperature of the warmest quarter was the most important bioclimatic variable determining its current range. High metabolic rate is a trait evolved to adapt to cold environment and thus cold temperature may not be a critical constrain to *L. variolus* comparing to the heat in summer [90].

Our models show the first reliable distribution maps of these two species and both species have potential distribution ranges beyond currently documented regions [43]. The logical next step is to conduct sampling efforts in typical habitats associated with these species in our predicted range to further validate the models. However, detection of range expansion of low density cryptic species such as *L. variolus* and *S. niger* would be the equivalent of searching for a needle in a hay stack for a small group of experts across such a wide region. Thus, we propose to call on citizen scientists by launching a monitoring project through a platform such as Bug-guide and iNaturalist to produce a large-scale sampling effort. This would represent a rapid, low cost, highly efficient, and innovative way to test these large scale predictive models. Although the risk of being bitten by northern black-widows is low, such a species monitoring approach mobilizing public will include safety guidelines to prevent health issues from the participants.

On the other hand, local health authorities should be informed of the documented presence of *L. variolus* north of its previously known range to have appropriate material and response protocols in case any citizen is bitten by this species, known to cohabit with humans in and around their buildings. For *S. niger*, new information about its potential range can serve as important guidelines to authorities and stakeholders working on conservation efforts for this species at various governmental levels where it is at risk.

Our distribution models not only increase our understanding of the current distribution of these two poorly documented spider species, they also provide guidance for corresponding public health and conservation management strategies. More importantly, we must emphasize that data collected from citizen science initiatives and other online scientific open-sources, with the incorporation of observer errors and spatial sampling bias parameters, enabled us to produce these reliable distribution models. Even if online information needs careful vetting before use, we show that citizen science initiatives can provide valuable occurrence data even for under-sampled rare species and habitat specialists. Open-access and digitalized data from museums and CS platforms will likely become an important and convenient data source for natural science research.

Supporting information

S1 File. *Latrodectus variolus* record. This file contains occurrence records of *Latrodectus variolus* collected and used in this research.

(XLS)

S2 File. *Sphodros niger* record. This file contains occurrence records of *Sphodros niger* collected and used in this research.

(XLS)

Acknowledgments

We sincerely thank all the citizen scientists who generously contribute their time and efforts in all the natural science open-access platforms. Without them and without these open-access platforms valuable quality data would not have been sufficient to make this project a reality. We would also like to thank Margot Charette for helping with collecting occurrence data of both spider species, Robb Bennett for kindly sharing his expertise on *Sphodros niger*, André-Philippe Drapeau Picard and Kent Mcfarland for their comments on previous versions of the manuscript.

Author Contributions

Conceptualization: Yifu Wang, Christopher Buddle, Maxim Larrivée.

Data curation: Yifu Wang, Maxim Larrivée.

Formal analysis: Yifu Wang, Nicolas Casajus, Maxim Larrivée.

Funding acquisition: Maxim Larrivée.

Investigation: Yifu Wang, Maxim Larrivée.

Methodology: Nicolas Casajus, Dominique Berteaux, Maxim Larrivée.

Project administration: Christopher Buddle, Maxim Larrivée.

Resources: Christopher Buddle, Dominique Berteaux, Maxim Larrivée.

Software: Nicolas Casajus, Maxim Larrivée.

Supervision: Christopher Buddle, Dominique Berteaux, Maxim Larrivée.

Validation: Maxim Larrivée.

Visualization: Nicolas Casajus.

Writing – original draft: Yifu Wang.

Writing – review & editing: Yifu Wang, Christopher Buddle, Maxim Larrivée.

References

1. Küper W, Sommer JH, Lovett JC, Barthlott W. Deficiency in African plant distribution data—missing pieces of the puzzle. *Botanical Journal of the Linnean Society*. 2006; 150(3):355–68.
2. Cao Y, DeWalt RE, Robinson JL, Tweddale T, Hinz L, Pessino M. Using Maxent to model the historic distributions of stonefly species in Illinois streams: the effects of regularization and threshold selections. *Ecological Modelling*. 2013; 259(0):30–9.
3. Newbold T. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography*. 2010; 34(1):3–22.
4. Breed GA, Stichter S, Crone EE. Climate-driven changes in northeastern US butterfly communities. *Nature Climate Change*. 2013; 3(2):142–5.

5. Devictor V, van Swaay C, Brereton T, Brotons L, Chamberlain D, Heliola J, et al. Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change*. 2012; 2(2):121–4.
6. Kullenberg C, Kasperowski D. What is citizen science? A scientometric meta-analysis. *PLOS ONE*. 2016; 11(1):e0147152. <https://doi.org/10.1371/journal.pone.0147152> PMID: 26766577
7. Bugguide. Bugguide 2015 [<http://bugguide.net/node/view/15740>].
8. GBIF. Global Biodiversity Information Facility (GBIF) 2015 [<http://www.gbif.org/>].
9. eBird. eBird: An online database of bird distribution and abundance Ithaca, New York 2015 [<http://www.ebird.org>].
10. eButterfly. eButterfly 2015 [<http://www.e-butterfly.org/#/>].
11. Encyclopedia of Life. Encyclopedia of Life 2015 [<http://eol.org/>].
12. iNaturalist. iNaturalist.org: A community for naturalists 2015 [<http://www.inaturalist.org/>].
13. Baboon Spider Atlas. Baboon Spider Atlas 2017 [<http://www.baboonspideratlas.co.za/>].
14. Donnelly A, Crowe O, Regan E, Begley S, Caffarra A. The role of citizen science in monitoring biodiversity in Ireland. *Int J Biometeorol*. 2014; 58(6):1237–49. <https://doi.org/10.1007/s00484-013-0717-0> PMID: 23975107
15. Gallo T, Waitt D. Creating a successful citizen science model to detect and report invasive species. *BioScience*. 2011; 61(6):459–65.
16. Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*. 2007; 34(1):102–17.
17. Kerr JT, Pindar A, Galpern P, Packer L, Potts SG, Roberts SM, et al. Climate change impacts on bumblebees converge across continents. *Science*. 2015; 349(6244):177–80. <https://doi.org/10.1126/science.aaa7031> PMID: 26160945
18. Parmesan C. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*. 2006; 37:637–69.
19. Porretta D, Mastrantonio V, Amendolia S, Gaiarsa S, Epis S, Genchi C, et al. Effects of global changes on the climatic niche of the tick *Ixodes ricinus* inferred by species distribution modelling. *Parasites Vectors*. 2013; 6(1):1–8.
20. Alto BW, Juliano SA. Precipitation and temperature effects on populations of *Aedes albopictus* (Diptera: Culicidae): implications for range expansion. *Journal of Medical Entomology*. 2001; 38(5):646–56. PMID: 11580037
21. Bale JS, Masters GJ, Hodkinson ID, Awmack C, Bezemer TM, Brown VK, et al. Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. *Global Change Biology*. 2002; 8(1):1–16.
22. Taucare-Ríos A, Bizama G, Bustamante RO. Using Global and Regional Species Distribution Models (SDM) to Infer the Invasive Stage of *Latrodectus geometricus* (Araneae: Theridiidae) in the Americas. *Environmental Entomology*. 2016; 45(6):1379–85. <https://doi.org/10.1093/ee/nvw118> PMID: 28028084
23. Vink CJ, Derraiik JGB, Phillips CB, Sirvid PJ. The invasive Australian redback spider, *Latrodectus hasseltii* Thorell 1870 (Araneae: Theridiidae): current and potential distributions, and likely impacts. *Biological Invasions*. 2011; 13(4):1003–19.
24. Năpăruș M, Kuntner M. A GIS Model Predicting Potential Distributions of a Lineage: A Test Case on Hermit Spiders (Nephilidae: Nephilengys). *PLOS ONE*. 2012; 7(1):e30047. <https://doi.org/10.1371/journal.pone.0030047> PMID: 22238692
25. Jiménez-Valverde A, Decae AE, Arnedo MA. Environmental suitability of new reported localities of the funnelweb spider *Macrothele calpeiana*: an assessment using potential distribution modelling with presence-only techniques. *Journal of Biogeography*. 2011; 38(6):1213–23.
26. Cramer KL, Maywright AV. Cold temperature tolerance and distribution of the brown recluse spider *Loxosceles reclusa* (Araneae, Sicariidae) in Illinois. *Journal of Arachnology*. 2008; 36(1):136–9.
27. Ammunt TA, Kaukoranta T, Saikkonen K, Repo T, Klemola T. Invading and resident defoliators in a changing climate: cold tolerance and predictions concerning extreme winter cold as a range-limiting factor. *Ecological Entomology*. 2012; 37(3):212–20.
28. Karban R, Strauss SY. Physiological tolerance, climate change, and a northward range shift in the spittlebug, *Philaenus spumarius*. *EEN Ecological Entomology*. 2004; 29(2):251–4.
29. Moriyama M, Numata H. Comparison of cold tolerance in eggs of two cicadas, *Cryptotympana facialis* and *Graptopsaltria nigrofuscata*, in relation to climate warming. *ENS Entomological Science*. 2009; 12(2):162–70.

30. Woinarski JCZ, Cullen JM. Distribution of invertebrates on foliage in forests of south-eastern Australia. *Australian Journal of Ecology*. 1984; 9(3):207–32.
31. Elith J, Graham CH, Anderson RP, Dudik M, et al. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*. 2006; 29(2):129–51.
32. Elith J, Kearney M, Phillips S. The art of modelling range-shifting species. *Methods in Ecology and Evolution*. 2010; 1(4):330–42.
33. Khatchikian C, Sangermano F, Kendell D, Livdahl T. Evaluation of species distribution model algorithms for fine-scale container-breeding mosquito risk prediction. *Medical and Veterinary Entomology*. 2011; 25(3):268–75. <https://doi.org/10.1111/j.1365-2915.2010.00935.x> PMID: 21198711
34. Hernandez PA, Graham CH, Master LL, Albert DL. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*. 2006; 29(5):773–85.
35. Kramer-Schadt S, Niedballa J, Pilgrim JD, Schröder B, Lindenborn J, Reinfelder V, et al. The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions*. 2013; 19(11):1366–79.
36. Hernandez PA, Franke I, Herzog SK, Pacheco V, Paniagua L, Quintana HL, et al. Predicting species distributions in poorly-studied landscapes. *Biodiversity and Conservation*. 2008; 17(6):1353–66.
37. Cuyckens GAE, Morales MM, Tognelli MF. Assessing the distribution of a vulnerable felid species: threats from human land use and climate change to the kodkod *Leopardus guigna*. *Oryx*. 2014; First-View:1–8.
38. Bergström U, Sundblad G, Downie A-L, Snickars M, Boström C, Lindegarth M. Evaluating eutrophication management scenarios in the Baltic Sea using species distribution modelling. *Journal of Applied Ecology*. 2013; 50(3):680–90.
39. Hu J, Liu Y. Unveiling the conservation biogeography of a data-deficient endangered bird species under climate change. *PLoS ONE*. 2014; 9(1):e84529. <https://doi.org/10.1371/journal.pone.0084529> PMID: 24404169
40. McCrone JD, Levi HW. North American widow spiders of the *Latrodectus curacaviensis* group (Araneae, Theridiidae). *Psyche*. 1964; 71(1):12–27.
41. Wilson LF. The Northern Widow Spider, *Latrodectus Variolus* (Araneae: Theridiidae), in Michigan. *The Great Lakes Entomologist*. 1967; 1(5):147–53.
42. Maretić Z. Latrodectism: variations in clinical manifestations provoked by *Latrodectus* species of spiders. *Toxicon*. 1983; 21(4):457–66. PMID: 6353667
43. NatureServe. NatureServe Explorer: An online encyclopedia of life Arlington, Virginia: Version 7.1. NatureServe; 2015 [<http://explorer.natureserve.org>].
44. Beatty JA. Web structure and burrow location of *Sphodros niger* (Hentz) (Araneae, Atypidae). *Journal of Arachnology*. 1986; 14(1):130–2.
45. McCook HC. Nesting habits of the American purseweb spider. *Proceedings of the Academy of Natural Sciences of Philadelphia*. 1888; 40:203–20.
46. Morrow W. A Range Extension of the Purseweb Spider *Sphodros Rufipes* in Eastern Kansas (Araneae, Atypidae). *The Journal of Arachnology*. 1986; 14(1):119–21.
47. Delaney DG, Sperling CD, Adams CS, Leung B. Marine invasive species: validation of citizen science and implications for national monitoring networks. *Biological Invasions*. 2008; 10(1):117–28.
48. Larrivé M, Kerr J. Eastern Canadian butterfly range expansions. *Bulletin of the Entomological Society of Canada*. 2012; 44(3):133–7.
49. Jensen GL, Lanier W, Seibert CE. Spider identification and management. Montana: Montana State University Extension Service. 2005.
50. Cutler B, Salsbury G, Guarisco H, Liggett C. The purse-web spiders of Kansas (Araneae: Atypidae). *Transactions of the Kansas Academy of Science*. 2004; 107(1 & 2):101–4.
51. Gertsch WJ, Platnick NI. A revision of the American spiders of the family Atypidae (Araneae, Mygalomorphae). *American Museum novitates*; no. 2704. 1980.
52. Coyle FA, Shear WA. Observations on the natural history of *Sphodros abboti* and *Sphodros rufipes* (Araneae, Atypidae), with evidence for a contact sex pheromone. *Journal of Arachnology*. 1981:317–26.
53. Wiczorek J, Bloom D, Constable H, Fang J, Koo M, Spencer C, et al. Georeferencing quick reference guide. 2012.
54. McKenney DW, Hutchinson MF, Papadopol P, Lawrence K, Pedlar J, Campbell K, et al. Customized spatial climate models for North America. *Bulletin of the American Meteorological Society*. 2011; 92(12):1611–22.

55. McKenney DW, Pedlar JH, Papadopol P, Hutchinson MF. The development of 1901–2000 historical monthly climate models for Canada and the United States. *Agricultural and Forest Meteorology*. 2006; 138(1–4):69–81.
56. Harte J. *Maximum entropy and ecology: a theory of abundance, distribution, and energetics*. Oxford; New York: Oxford University Press; 2011.
57. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*. 2006; 190(3–4):231–59.
58. Li X, Wang Y. Applying various algorithms for species distribution modelling. *Integrative Zoology*. 2013; 8(2):124–35. <https://doi.org/10.1111/1749-4877.12000> PMID: 23731809
59. Rubio GD, Rodrigues ENL, Acosta LE. Description of the male of the spider *Dubiaranea difficilis* (Araneae: Linyphiidae), with new records and modeling of its potential geographic distribution. *Zootaxa*. 2010(2405):55–62.
60. Adams-Hosking C, Grantham HS, Rhodes JR, McAlpine C, Moss PT. Modelling climate-change-induced shifts in the distribution of the koala. *Wildlife Research*. 2011; 38(2):122–30.
61. Leroux SJ, Larrivée M, Boucher-Lalonde V, Hurford A, Zuloaga J, Kerr JT, et al. Mechanistic models for the spatial spread of species under climate change. *Ecological Applications*. 2013; 23(4):815–28. PMID: 23865232
62. McFarland KP, Rimmer CC, Goetz JE, Aubry Y, Wunderle JM Jr., Sutton A, et al. A winter distribution model for Bicknell's Thrush (*Catharus bicknelli*), a conservation tool for a threatened migratory songbird. *PLOS ONE*. 2013; 8(1):e53986. <https://doi.org/10.1371/journal.pone.0053986> PMID: 23326554
63. Allouche O, Tsoar A, Kadmon R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*. 2006; 43(6):1223–32.
64. Landis JR, Koch GG. The measurement of observer agreement for categorical Data. *Biometrics*. 1977; 33(1):159–74. PMID: 843571
65. Liu C, White M, Newell G. Measuring and comparing the accuracy of species distribution models with presence–absence data. *Ecography*. 2011; 34(2):232–43.
66. Swets JA. Measuring the accuracy of diagnostic systems. *Science*. 1988; 240(4857):1285–93. PMID: 3287615
67. Phillips SJ, Dudík M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*. 2008; 31(2):161–75.
68. Boria RA, Olson LE, Goodman SM, Anderson RP. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*. 2014; 275(0):73–7.
69. Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, et al. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*. 2009; 19(1):181–97. PMID: 19323182
70. Warren DL, Wright AN, Seifert SN, Shaffer HB. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. *Diversity and Distributions*. 2014; 20(3):334–43.
71. Clements GR, Rayan DM, Aziz SA, Kawanishi K, Traeholt C, Magintan D, et al. Predicting the distribution of the Asian tapir in Peninsular Malaysia using maximum entropy modeling. *Integrative Zoology*. 2012; 7(4):400–6. <https://doi.org/10.1111/j.1749-4877.2012.00314.x> PMID: 23253371
72. Syfert MM, Smith MJ, Coomes DA. The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. *PLoS ONE*. 2013; 8(2):e55158. <https://doi.org/10.1371/journal.pone.0055158> PMID: 23457462
73. Quillfeldt P, Masello JF, Navarro J, Phillips RA. Year-round distribution suggests spatial segregation of two small petrel species in the South Atlantic. *Journal of Biogeography*. 2013; 40(3):430–41.
74. North GR. More evidence for anthropogenic influence on climate change. *Proceedings of the National Academy of Sciences*. 2013; 110(43):17169–70.
75. Proosdij ASJ, Sosef MSM, Wieringa JJ, Raes N. Minimum required number of specimen records to develop accurate species distribution models. *Ecography*. 2016; 39(6):542–52.
76. Dickinson JL, Zuckerberg B, Bonter DN. Citizen science as an ecological research tool: challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics*. 2010; 41(1):149–72.
77. Barghusen LE, Claussen DL, Anderson MS, Bailer AJ. The effects of temperature on the web-building behaviour of the common house spider, *Achaearanea tepidariorum*. *Functional Ecology*. 1997; 11(1):4–10.
78. Bonte D, Travis JMJ, De Clercq N, Zwertvaegher I, Lens L. Thermal conditions during juvenile development affect adult dispersal in a spider. *Proceedings of the National Academy of Sciences*. 2008; 105(44):17000–5.

79. Kiss B, Samu F. Comparison of autumn and winter development of two wolf spider species (*Pardosa*, Lycosidae, Araneae) having different life history patterns. *Journal of Arachnology*. 2002; 30(2):409–15.
80. Pruitt JN, Demes KW, Dittrich-Reed DR. Temperature mediates shifts in individual aggressiveness, activity level, and social behavior in a spider. *Ethology*. 2011; 117(4):318–25.
81. Dube D-e. Black Widow spider shocks Barrhaven family with visit. *Ottawa Sun*. 2015.
82. Poporat E. North Black Widow observation. In: Larrivée M, editor. 2014.
83. IPCC. Climate change 2014: synthesis report. Contribution of working groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva. Switzerland; 2014.
84. Saupé EE, Papes M, Selden PA, Vetter RS. Tracking a Medically Important Spider: Climate Change, Ecological Niche Modeling, and the Brown Recluse (*Loxosceles reclusa*). *PLOS ONE*. 2011; 6(3): e17731. <https://doi.org/10.1371/journal.pone.0017731> PMID: 21464985
85. Finkbeiner SD, Reed RD, Dirig R, Losey JE. The role of environmental factors in the northeastern range expansion of *Papilio cresphontes* Cramer (Papilionidae). *Journal of the Lepidopterists Society*. 2011; 65(2):119–25.
86. Greenstone MH. Ballooning frequency and habitat predictability in two wolf spider species (Lycosidae: *Pardosa*). *The Florida Entomologist*. 1982; 65(1):83–9.
87. Larrivée M, Buddle CM. Ballooning propensity of canopy and understorey spiders in a mature temperate hardwood forest. *Ecological Entomology*. 2011; 36(2):144–51.
88. Coyle FA, Greenstone MH, Hultsch A-L, Morgan CE. Ballooning Mygalomorphs: Estimates of the Masses of *Sphodros* and *Ummidia* Ballooners (Araneae: Atypidae, Ctenizidae). *The Journal of Arachnology*. 1985; 13(3):291–6.
89. Anderson JF. Comparative energetics of comb-footed spiders (Araneae: Theridiidae). *Comparative Biochemistry and Physiology Part A: Physiology*. 1994; 109(1):181–9.
90. Addo-Bediako A, Chown SL, Gaston KJ. Metabolic cold adaptation in insects: a large-scale perspective. *Functional Ecology*. 2002; 16(3):332–8.