Progressive deterioration of pollination service detected in a 17-year study vanishes in a 26-year study

James D. Thomson[1,2]

[1]Department of Ecology and Evolutionary Biology, University of Toronto, 25 Harbord St., Toronto, ON M5S 3G5, Canada; [2]Rocky Mountain Biological Laboratory, Crested Butte, CO 81224, USA

Author for correspondence:
James D. Thomson
Tel: +1 416 978 3527
Email: james.thomson@utoronto.ca

Received: *12 March 2019*
Accepted: *17 July 2019*

New Phytologist (2019)
doi: 10.1111/nph.16078

Key words: *Erythronium grandiflorum*, phenological dislocation, pollen supplementation, pollination deficit, pollination limitation, pollinator decline.

Summary

- Widespread reports of declining populations of pollinators have raised concerns that plant populations may be incurring increasing shortfalls in pollination, but few studies have measured pollination deficits over enough seasons to detect such changes. I have conducted pollen-supplementation experiments in a wild population of the glacier lily (*Erythronium grandiflorum*, Liliaceae) from 1993 to 2018.
- Pollination deficits were estimated by comparing the fruit set of hand-pollinated, single-flowered plants to that of open-pollinated controls. For a subset of years, seed set data were also available.
- A previous publication reported a significant deterioration of pollination in this population from 1993 to 2009, and suggested phenological dislocation as a possible cause. That deterioration is no longer evident in the longer-term data set.
- Very long time series may be necessary to detect temporal trends in pollination service. This population consistently experiences stronger pollination deficits before its flowering peak than after. This heterogeneity suggests caution in characterizing a population as pollination-limited or not, even within a single season.

Introduction

Reports of declining populations or shrinking distributions of flower-feeding animals (e.g. National Research Council, 2006, Williams & Osborne, 2009; Potts *et al.*, 2010) have prompted predictions of declining pollination service. Although such predictions are plausible, the most frequently cited evidence is indirect: some plant populations have declined in tandem with some groups of flower visitors. As Biesmeijer *et al.* (2006) conclude in one influential study, ‘We cannot tell from these data whether the decline of the plants precedes the loss of the associated pollinators, whether the decline of the pollinators leads to the loss of reproductive function and then to the decline of the plants, or indeed whether the plants and their pollinators are both responding to some other factor.’

Although direct evidence of deteriorating pollination service – in the form of experimental measurements over time – remains scant, the idea has engaged researchers. As of 27 June 2019, Web of Science returned 1662 references for the phrase ‘pollination decline’, and another 87 for the more strident phrase ‘pollination crisis’. Google searches for these phrases produce hundreds of thousands of hits, indicating that the laity has joined the discussion. Much of this attention concerns honey bees and agricultural crops (for which Martin (2015) provides a measured review), but wild plants are also of concern.

We cannot consider temporal changes in pollination success without a reliable way to measure pollination success in the first place. Simply measuring fruit or seed production will seldom suffice, because those processes are heavily influenced by uncontrolled environmental variation across space and seasons. To control for such variation, a common approach is to measure pollination deficits experimentally within local study populations, comparing an open-pollinated *control* group of plants to a *supplemented* group in which all flowers have additionally been pollinated by hand. The idealized concept is that the supplemented plants will have received ‘complete’ pollination, and therefore can serve as a standardized benchmark for comparing the performance of the unmanipulated group. Its simplicity makes this experimental design attractive, and pollination ecologists have used it to assess pollination sufficiency in hundreds of natural populations. Reviews of this literature (Burd, 1994; Knight *et al.*, 2005, 2006; García-Camacho & Totland, 2009; Rosenheim *et al.*, 2014) confirm that pollination deficits, namely cases of ‘pollination limitation’ of seed or fruit production, are very common, although not universal.

Despite their popularity and apparent simplicity, supplemental-pollination experiments suffer from biases. Because the added pollen is usually outcross pollen, it will be of higher quality than the mixture of self and outcross pollen that cosexual plants will usually receive from animal visits. By thus setting a benchmark for pollination quality that is unachievable in nature, such experiments are biased toward exaggerating pollination deficits (Thomson, 2001; Aizen & Harder, 2007). A second bias, in the opposite direction, is driven by the tendency of researchers to choose study sites with dense, healthy stands of plants. Those plants are likely to receive better pollination service than members of sparse, peripheral populations (e.g. Ågren, 1996; reviewed by Wilson *et al*., 1994). Knight *et al.* (2006) further point out that pollinating only a subset of a plant's flowers, or not applying treatments for the life of a plant, will also bias results toward pollination limitation. Bos *et al.* (2007) argue that seed set data are more informative than the more commonly recorded fruit set. Depending on how they are conducted, therefore, supplementation experiments are variously susceptible to over- or under-estimation of the 'true' (i.e. evolutionarily relevant) magnitude of a pollination shortfall in any particular circumstance. Despite these ambiguities, they can provide a robust and repeatable method for answering the narrow question of whether pollination service is getting better or worse in a single, stable population that has been studied over many seasons.

I previously reported (Thomson, 2010) a progressive increase in pollination deficits over a 17-yr study of a single population of a long-lived perennial geophyte, the glacier lily *Erythronium grandiflorum*. Since then, I have continued to perform annual supplementation experiments at the same site with the same methodology. This brief update covers the 26-yr period from 1993 to 2018. The primary question remains the same: Is pollination service improving, deteriorating, or remaining stable? Notably, the additional data impel a different conclusion.

Materials and Methods

Because the site and the experimental procedures have not changed from the detailed descriptions published earlier (Thomson, 2010), I provide only a brief sketch here. The 2010 paper also gives further information and references on the ecology and reproductive biology of the study species. Located at 3170 m elevation in the Gunnison National Forest of the Colorado Rocky Mountains, the study site is a private inholding of about 3 ha. The vegetation is upland subalpine meadow with scattered spruce and fir trees. The long-lived perennial *Erythronium grandiflorum* Pursh (Liliaceae) ('glacier lily') is the first plant to bloom after snowmelt in late May to early June. Shoots often emerge through the last few centimeters of receding snow. Over 90% of plants produce a single flower per year. Large interannual variation in snow depth means that the earliest springs begin > 1 month earlier than the latest springs. Despite the variation in onset dates, the spatial progression of snowmelt across the site is highly consistent and predictable across years (Thomson, 2010; see Yamagishi *et al*., 2005). The flowering period varies with weather, but usually lasts about 4 wk. Bumble bee queens, the principal pollinators, begin emerging around the onset of flowering.

I first conducted supplemental-pollination experiments in 1993. After a few years of doing one or two experiments per season near the peak of bloom, in 1998 I settled on a fixed routine of three experiments (*early*, *middle* and *late*) per year. The *middle* treatment was intended to coincide with peak flowering, while the *early* and *late* treatments were intended to approximate the first and last deciles of flowering, respectively. The progression of bloom depended on weather, so the intervals between successive treatments varied by a day or two. I conducted pollinations only in fair weather. In the figures, data points from *early*, *middle* and *late* are shown in yellow, orange and red, respectively.

To standardize pollen freshness, I picked half-dehisced anthers from at least 20 plants near the center of the population and let them finish dehiscing in a polystyrene vial. Using surveyor's pin flags with their shafts painted to designate pollination treatments, I then marked 150 neighboring pairs of single-flowered plants, moving through the population in one direction so that results would not be dominated by local patchiness. To standardize receptivity, I chose recipient flowers with more than three anthers dehisced. If all six anthers had dehisced, the flowers had to look fresh, with no sign of bleaching or wilting. I matched paired flowers for age and stature, then used a Microbrush to stir up the donor pollen mixture and completely cover the stigma of the pair member chosen for pollination. Approximately 6 wk after pollination, I scored whether flowers set fruit. In most years after 1996, I harvested and pickled all fruits before they dehisced, so that seed set could be assessed by later dissections.

Flowers can fail for various causes other than inadequate pollination. Because *Erythronium* blooms so early, late snow or hail storms and hard frosts can destroy cohorts of flowers wholly or partially (see Inouye, 2000, 2008). Complete destruction makes it impossible to compute pollination deficits. Partial destruction allows computation, but reduces sample sizes. A few flowers are eaten shortly after pollination, typically by chipmunks; these leave a shriveled, thread-like scape, and are excluded from analysis. Later, some mature fruits are taken by deer, leaving behind a sturdy, straight, torn-off scape. These are counted as fruits.

Statistical analysis

Various indices of pollination limitation are available, but all depend on the difference between a response variable measured for the open control group and the supplementally pollinated group. In the earlier paper (Thomson, 2010) I deferred to a reviewer who preferred a log-odds ratio ($\log_e$(supp/control)) for its superior statistical properties. Here, I defer to a second reviewer who favored fractional measures as easier to interpret. For my fruit set data (Table 1), the two indices are so highly correlated (Pearson's $r = 0.986$, $n = 48$) that they clearly capture identical information.

My primary analysis concerns the response of fruit set only, ignoring seeds (Table 1). I first calculate the proportion of flowers that set fruit set in each treatment, then divide the fruit success for the supplemental group by the sum of the successes of the control group and the supplemented group (i.e. supp/(supp + control)). This index of pollination limitation ranges from 0 to 1, with values above 0.5 indicating pollination deficits. It is effectively equivalent to Larson & Barrett's (2000) index (1 − (control/supp)), but remains defined in the unlikely event that supp = 0. I use simple linear regressions of this index on years to test the hypothesis that pollination service has

Table 1 Fruit set data, from 1993 to 2018, in phenological cohorts of supplementally (Supp.) pollinated flowers and matched open-pollinated controls in a population of the wild lily *Erythronium grandiflorum*.

Year	Date class	Treatment	Total sample	Fruits	Failed fruits	Fruit set	Fruit set limitation index	*P*-value
1993	Middle	Supp.	150	5	144	0.034	0.386	0.402
		Control	150	8	142	0.053		
1994	Middle	Supp.	150	121	17	0.877	0.498	0.888
		Control	150	120	16	0.882		
1995	Middle	Supp.	150	85	65	0.567	0.521	0.417
		Control	150	78	72	0.520		
1997	Middle	Supp.	175	105	65	0.618	**0.582**	0.001
		Control	180	79	99	0.444		
	Late	Supp.	50	30	19	0.612	0.586	0.082
		Control	50	19	25	0.432		
1998	Early	Supp.	150	0	148	0.000	Undefined	NS
		Control	150	0	143	0.000		
	Middle	Supp.	149	0	149	0.000	Undefined	NS
		Control	151	0	151	0.000		
	Late	Supp.	149	121	28	0.812	0.503	0.814
		Control	151	121	30	0.801		
1999	Early	Supp.	150	39	111	0.260	**0.886**	<0.0001
		Control	150	5	145	0.033		
	Middle	Supp.	150	115	35	0.767	**0.714**	<0.0001
		Control	150	46	104	0.307		
	Late	Supp.	150	82	68	0.547	**0.617**	0.000
		Control	150	51	99	0.340		
2003	Early	Supp.	147	114	30	0.792	**0.736**	<0.0001
		Control	149	42	106	0.284		
	Middle	Supp.	118	64	53	0.547	**0.660**	<0.0001
		Control	124	35	89	0.282		
	Late	Supp.	150	69	81	0.460	**0.620**	0.001
		Control	149	42	107	0.282		
2005	Early	Supp.	144	84	50	0.627	**0.558**	0.029
		Control	149	70	71	0.496		
	Middle	Supp.	106	85	17	0.833	**0.611**	<0.0001
		Control	103	53	47	0.530		
	Late	Supp.	159	128	19	0.871	**0.566**	<0.0001
		Control	150	95	47	0.669		
2006	Early	Supp.	148	81	60	0.574	**0.673**	<0.0001
		Control	149	39	101	0.279		
	Middle	Supp.	145	46	95	0.326	0.563	0.178
		Control	145	36	106	0.254		
	Late	Supp.	137	92	43	0.681	0.535	0.131
		Control	134	77	53	0.592		
2008	Early	Supp.	144	89	44	0.669	**0.643**	<0.0001
		Control	142	49	83	0.371		
	Middle	Supp.	148	100	39	0.719	**0.676**	<0.0001
		Control	149	48	91	0.345		
	Late	Supp.	145	93	46	0.669	**0.649**	<0.0001
		Control	148	50	88	0.362		
2009	Early	Supp.	145	10	135	0.069	**0.909**	0.006
		Control	146	1	143	0.007		
	Middle	Supp.	132	27	104	0.206	**0.904**	<0.0001
		Control	139	3	134	0.022		
	Late	Supp.	145	75	70	0.517	**0.640**	0.000
		Control	148	43	105	0.291		
2010	Early	Supp.	147	13	131	0.090	0.655	0.154
		Control	148	7	140	0.048		
	Middle	Supp.	148	6	142	0.041	0.425	0.566
		Control	146	8	138	0.055		
	Late	Supp.	147	23	124	0.156	0.568	0.354
		Control	143	17	126	0.119		

Table 1. (Continued)

Year	Date class	Treatment	Total sample	Fruits	Failed fruits	Fruit set	Fruit set limitation index	*P*-value
2011	Early	Supp.	130	89	37	0.706	**0.610**	<0.001
		Control	142	64	78	0.451		
	Middle	Supp.	136	102	16	0.864	**0.568**	<0.001
		Control	144	77	40	0.658		
	Late	Supp.	142	92	24	0.793	0.523	0.210
		Control	142	89	34	0.724		
2012	Early	Supp.	No flowers available, hail					
		Control						
	Middle	Supp.	No flowers available, hail					
		Control						
	Very late	Supp.	149	61	86	0.415	0.565	0.091
		Control	145	46	98	0.319		
2013	Early	Supp.	150	128	22	0.853	**0.759**	<0.001
		Control	149	40	108	0.270		
	Middle	Supp.	142	112	29	0.794	**0.641**	<0.001
		Control	149	65	81	0.445		
	Late	Supp.	150	120	28	0.811	0.492	0.543
		Control	143	119	23	0.838		
2014	Early	Supp.	Not done, all flowers killed by hail					
		Control						
	Middle	Supp.	147	48	97	0.331	**0.603**	0.033
		Control	145	31	111	0.218		
	Late	Supp.	152	125	27	0.822	**0.542**	0.010
		Control	144	100	44	0.694		
2015	Early	Supp.	141	62	79	0.440	**0.713**	<0.001
		Control	142	25	116	0.177		
	Middle	Supp.	142	113	29	0.796	**0.678**	<0.001
		Control	151	57	94	0.377		
	Late	Supp.	147	136	11	0.925	0.502	0.775
		Control	143	131	12	0.916		
2016	Early	Supp.	146	99	45	0.688	**0.584**	<0.001
		Control	141	69	72	0.489		
	Middle	Supp.	140	80	60	0.571	**0.573**	0.014
		Control	150	64	86	0.427		
	Late	Supp.	143	80	63	0.559	0.490	0.683
		Control	144	84	60	0.583		
2017	Early	Supp.	145	123	22	0.848	**0.534**	0.022
		Control	143	105	37	0.739		
	Middle	Supp.	142	90	52	0.634	0.489	0.616
		Control	145	96	49	0.662		
	Late	Supp.	148	124	24	0.838	0.514	0.309
		Control	149	118	31	0.792		
2018	Early	Supp.	73	48	15	0.762	**0.597**	0.003
		Control	74	36	34	0.514		
	Middle	Supp.	74	46	24	0.657	0.507	0.820
		Control	75	46	26	0.639		
	Late	Supp.	75	59	12	0.831	0.484	0.376
		Control	70	54	7	0.885		

In 1993–1995, a single experiment was done near the peak date of flowering. In 1997, a second, smaller experiment was added after the peak. In subsequent years, three experiments of *c*. 300 flowers were done, before, at and after the peak ('early, middle and late'). 'Failed fruits' include those lost to frost and to abortion. Fruit set is calculated as the number of fruits divided by the total number of successful and failed fruits. The index of 'Fruit set limitation' equals the fruit set for the supplemented flowers divided by the sum of the fruit sets for supplemented and control flowers; values of 0.5 indicate equal success of supplemented and control flowers, whereas higher values indicate greater success of supplemented flowers (i.e. pollination limitation), with significant values in bold type.

improved or deteriorated over the study period. Such regressions are simple to interpret within each of the *early*, *middle* and *late* experimental groups. In the earlier study (see Thomson, 2010, fig. 5), I wished to combine the data from all three groups, to obtain an overall test with the greatest possible number of data points. That required an *ad hoc* adjustment of the data to equalize mean pollination deficits among the three groups. With more data now available, clear conclusions can be reached without combining the three subgroups, so I omit the combination-and-adjustment procedure.

A secondary analysis considers seed-set data for the years in which those data are available. Fruits were harvested when mature seeds could be unambiguously distinguished from unfertilized or aborted ovules. First, I used seed counts (Table 2) to compute an alternative index of pollination limitation based on seeds/fruit rather than on fractional seed set. I examined the linear correlation between the seed-set index and the fruit-set index described above. Second, I computed a 'total', or seed-weighted, index of pollination limitation in which the fractional fruit set within a treatment is multiplied by the mean seeds/fruit for that treatment.

Simply regressing these indices against time gives each experiment equal weight. The sample sizes are large enough (typically 150 plants per treatment) to provide reliable estimates of response variables. In some years, however, numerous flowers were lost to weather. In a few cases, data points based on few flowers exert strong leverage on regression lines. In those cases, I calculate fits with and without the outliers, giving details in the figure captions. For the primary test of fruit-set limitation over time, I fitted an alternative, logistic regression analysis that obviates the sample-size question by treating individual flowers as binomial responses. This model is not presented here because of its complexity, but it confirms the simpler analyses. R code is available on request.

Results

Pollination limitation based on fruit set

Fig. 1 shows linear regressions of the pollination-deficit index across the years, separately for the *early*, *middle* and *late* subsets of experiments. In the earlier analysis of data from 1993 from 2009, all three temporal subsets had positive slopes indicating increasing deficits. When those three subsets were pooled to increase the sample size, after adjusting the data to equalize the means of the groups, the trend toward worsening pollination service became significant at $P = 0.03$. In the newly extended dataset, the *middle* group's slope continues to suggest a slight deterioration of pollination, but with a negligible R^2 of 0.028. The *early* and *late* group slopes have now shifted to suggest slight trends in the opposite direction, namely toward improved pollination, but none of the three regressions are statistically significant. Pooling the three subsets, as in Thomson (2010), would not alter the conclusion that this relationship is flat.

Although there is no directional change in pollination service, the data are not without pattern. Inspection of Table 1 indicates that pollination deficits (i.e. indices > 0.5, with significant χ^2 tests) remain common. Still, the magnitude of deficits varies across and within years. As Fig. 2 indicates, pollination deficits are higher for the early-blooming portion of the population. Table 1 shows that deficits often decline to statistical nonsignificance in the later period of bloom.

Pollination limitation based on seed set

It is possible that the fraction of flowers setting fruit and the numbers of seeds per fruit might respond differently to variation in pollination service. Such differences would complicate the interpretation of pollination deficits (see Bos *et al.*, 2007); essentially, one conceptual quantity would be measured by two operational definitions that yield different results. Table 2 presents seed-set data for the subset of years in which seeds were counted. That table also gives pollination-limitation indices for seeds/per fruit, calculated analogously to the fruit-set indices discussed above. The two indices are highly correlated across all experiments (Fig. 3), indicating that the two operational measures of pollination service are interchangeable in this population. Either fruit-set data alone or seed-set data alone can serve as reliable response variables.

Combining fruit and seed data

To explore the joint response of fruit and seed production to supplemental pollination, I refined the fruit-set index (Table 1, Fig. 1) by weighting the proportional fruit set term for each treatment by the mean number of seeds per fruit for that treatment (Table 2). This yields a 'total' index that more completely characterizes female reproductive success. Presenting this index graphically across years can potentially provide a more meaningful portrayal of temporal trends in pollination service (Fig. 4) than the fruit-set index alone (Fig. 1), although it comes at the cost of a smaller number of data points. Because it depends on the multiplication of two terms that are themselves highly correlated, the seed-weighted index does not present qualitatively new patterns; rather, it shows how the addition of seed-set data amplifies the responses shown by fruit-set data alone. Interestingly, all three temporal subsets now have slopes suggesting improvement in pollination service rather than deterioration, and those trends reach conventional statistical significance for the *early* and *late* groups, if not for the middle group. Given the small number of data points, it would be unwise to take these trends very seriously, but they certainly reinforce the primary message: the previously reported pattern of increasing deficits is no longer evident.

Discussion

This study has produced large data sets that could be used to examine various issues, but the primary scope of this paper is narrow by design. Rather than seeking to model the dependence of reproductive success on various factors, I seek only to extract indices of pollination service that can establish whether plants commonly suffer pollination deficits in this system (the answer is yes) and whether such deficits have intensified over recent years, as previously reported (the answer is no).

The evident deterioration of pollination service for the period 1993–2009 prompted speculation about possible causes. In the previous report (Thomson, 2010), I considered two plausible, nonexclusive explanations: bumble bee abundance or activity may have declined, or a phenological dislocation may have occurred (e.g. Memmott *et al.*, 2007; Hegland *et al.*, 2009; Forrest & Thomson, 2011), with flowering increasingly beginning before the queen bees emerged and began foraging. Lacking quantitative estimates of bee activity, I could not rigorously evaluate either hypothesis, but I argued that the phenological

Table 2 Relationship between fruit set (see Table 1) and seed set in pollen-supplementation experiments in a population of the wild lily *Erythronium grandiflorum*.

		Early supplementation experiments						Middle (peak) supplementation experiments						Late supplementation experiments					
Year	S = supplemented, C = control	Fractional fruit set	Mean seed number per fruit	Total seed production	Fractional fruit limitation index	Fractional seed limitation index	Total seed-weighted limitation index	Fractional fruit set	Mean seed number per fruit	Total seed production	Fractional fruit limitation index	Fractional seed limitation index	Total seed-weighted limitation index	Fractional fruit set	Mean seed number per fruit	Total seed production	Fractional fruit limitation index	Fractional seed limitation index	Total seed-weighted limitation index
1997	S							0.618	18.720	11.569	0.582	0.568	0.647						
	C							0.444	14.230	6.318									
1999	S	0.260	21.160	5.502	0.887	0.803	0.970	0.767	20.510	15.731	0.714	0.617	0.801	0.547	28.496	15.587	0.617	0.607	0.713
	C	0.033	5.200	0.172				0.307	12.730	3.908				0.340	18.435	6.268			
2003	S	0.792	31.470	24.924	0.736	0.648	0.837	0.547	31.400	17.176	0.660	0.615	0.756	0.460	23.696	10.900	0.620	0.518	0.637
	C	0.284	17.075	4.849				0.282	19.677	5.549				0.282	22.030	6.212			
2008	S	0.475	27.730	13.172	0.572	0.591	0.658	0.351	20.330	7.136	0.589	0.580	0.665	0.370	20.511	7.589	0.535	0.533	0.568
	C	0.356	19.230	6.846				0.245	14.700	3.602				0.322	17.960	5.783			
2011	S	0.706	21.559	15.221	0.610	0.515	0.624	0.864	27.099	23.414	0.568	0.529	0.596	0.793	27.923	22.143	0.523	0.496	0.518
	C	0.451	20.321	9.165				0.658	24.131	15.878				0.724	28.414	20.572			
2012	S													0.415	25.083	10.409	0.565	0.448	0.514
	C													0.319	30.889	9.854			
2014	S							0.331	19.409	6.424	0.603	0.555	0.655	0.822	27.248	22.398	0.542	0.521	0.563
	C							0.218	15.552	3.390				0.694	25.088	17.411			
2017	S	0.848	28.455	24.130	0.534	0.516	0.550	0.634	22.307	14.143	0.489	0.523	0.512	0.838	23.127	19.380	0.514	0.468	0.482
	C	0.739	26.710	19.739				0.662	20.374	13.488				0.792	26.277	20.811			

Seed data are not available for all experiments. Fractional fruit set = fruits/flowers; Total seed production = (mean seed number per fruit) × (number of fruits set); both Fractional limitation indices are calculated as the values for S/(S + C), so values above 0.5 indicate pollination limitation. The correlation between fruit and seed limitation indices is presented in Fig. 3.

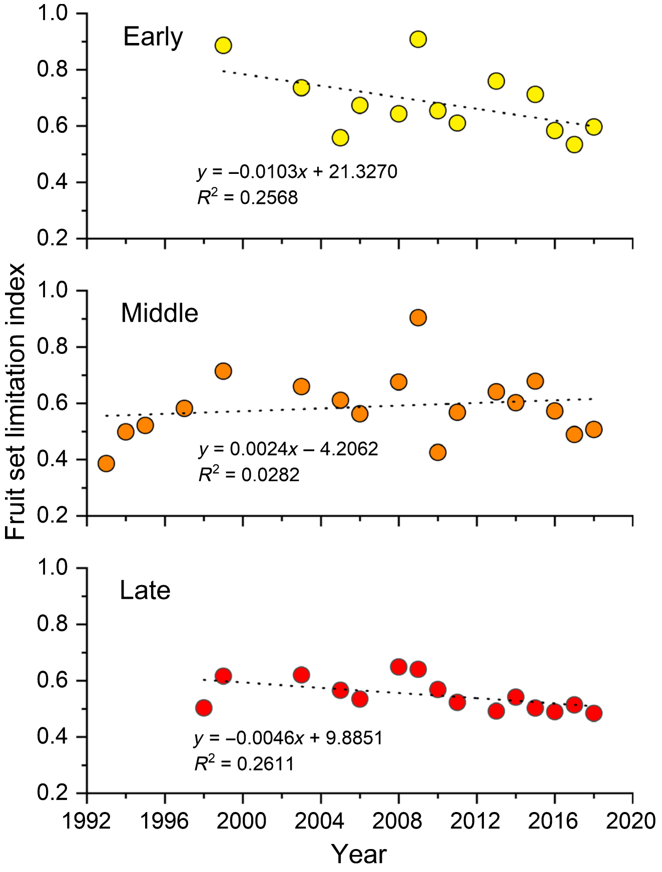

Fig. 1 Indices of pollination deficits, based on fruit sets in supplemental pollination experiments conducted on *Erythronium grandiflorum* lilies at Irwin, Colorado, from 1993 to 2018. Data are presented in Table 1. This fruit-set-based index is calculated as (fractional fruit set of supplemented plants/sum of fractional fruit sets for supplemented and control treatments), so a value of 0.5 indicates equal fruit set in the two treatments. Panels show phenological subsets of experiments (*early*, *middle* and *late*, indicated in all figures by yellow, orange and red color codes) that were conducted at three different times during each flowering season. Dotted lines show standard linear regressions. None of them are significant at $P < 0.05$; the slopes are presented only to show the trends in the data.

dislocation hypothesis was consistent with two observations: the observed trend of increasingly early snowmelt and correspondingly earlier flowering of *E. grandiflorum* over the years of observation (Miller-Rushing & Inouye, 2009), and the tendency of earlier flowers to incur large pollination deficits (Thomson, 2010). Now that the decline in pollination has vanished, questions about causative factors have simply lost meaning.

What explains the different messages conveyed by the 1993–2009 and the 1993–2018 data sets? Have important ecological factors changed? The pattern of deficit indices through time (Fig. 1) suggests not. Pollination deficits are frequent, especially in the *early* experiments, but their strength varies across years. It appears that the 1993–2009 period started, by chance, with some particularly good years for pollination and ended, also by chance, with some bad ones. The significant regression in the 2010 paper was simply a consequence of choosing an unfortunate window of time.

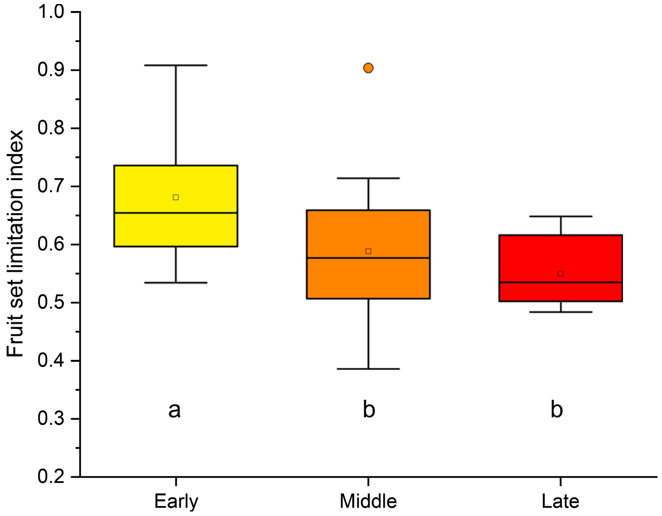

Fig. 2 Boxplots of indices of pollination deficits based on fruit set for the *early*, *middle* and *late* subsets of supplemental pollination experiments conducted on *Erythronium grandiflorum* lilies at Irwin, Colorado, from 1993 to 2018. An index value of 0.5 indicates equal fruit set in supplementally pollinated and open-pollinated plants. Colored boxes indicate interquartile ranges around the horizontal lines representing the medians. Outer whiskers denote 5th and 95th percentiles, and a single outlier is shown for the *middle* group. Small squares within the boxes denote means. Means (standard deviations) are 0.681 (0.117) for *early*, 0.589 (0.118) for *middle* and 0.550 (0.057) for *late*. Letters (a, b) under the boxes denote groups that differ significantly ($P < 0.04$) according to two-tailed *t*-tests for unequal variances. *Early* cohorts experience more pollination limitation compared with later cohorts.

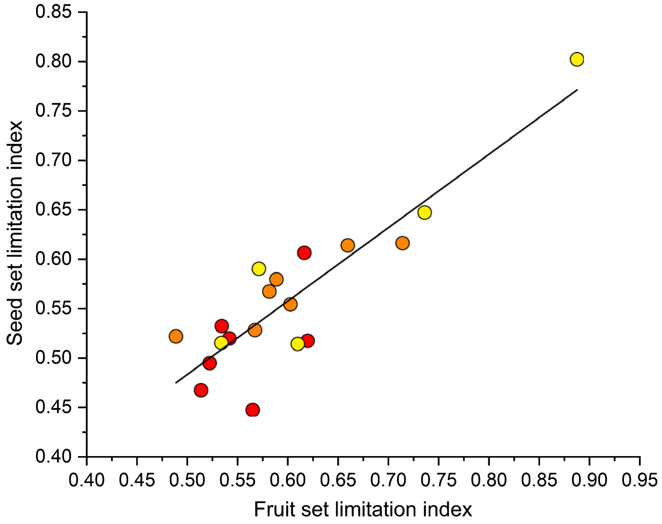

Fig. 3 Correlation between indices of pollination limitation based on fruit set and seed set, from the data presented in Table 2. For the full data set as shown, the Pearson correlation is 0.958 ($n = 19$); if the outlying point at the upper left is removed, the correlation drops to 0.942 ($n = 18$); both relationships are highly significant ($P < 10^{-8}$). Dropping the outlier changes the slope by only a small amount, from 0.71 to 0.61.

Therefore, it no longer seems likely that the system is being stressed by progressive, climate-driven phenological dislocation. This result is unsurprising, given that we would expect high-altitude organisms to be adapted to endure unpredictable weather.

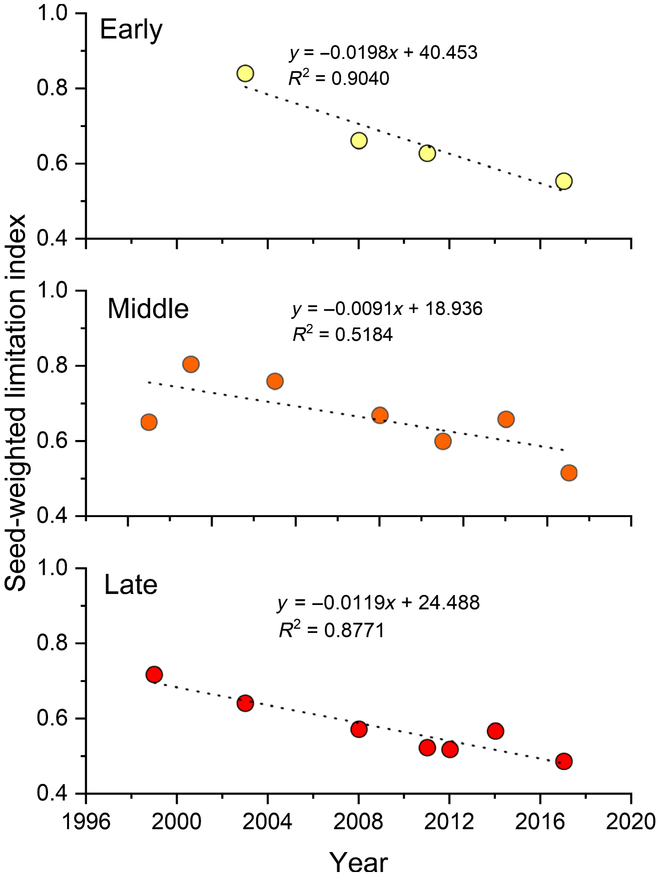

Fig. 4 Indices of pollination deficits, based on fruit and seed sets in supplemental pollination experiments conducted on *Erythronium grandiflorum* lilies at Irwin, Colorado, from 1997 to 2017, as presented in Table 2. This seed-weighted limitation index is calculated as in Fig. 1, but the fractional fruit-set values are now weighted by the mean number of seeds produced by the fruits of each treatment. Panels show phenological subsets of experiments (*early*, *middle* and *late*) that were conducted at three different times during each flowering season. The outlying datum for 1999 *early* has been removed because the seed counts were based on only five control fruits (see Table 1). Dotted lines show standard linear regressions. None of them have positive slopes significant at $P < 0.05$. Correlations reach conventional significance for the *early* ($P = 0.49$) and *late* ($P = 0.002$) subsets, but not for the *middle* subset ($P = 0.068$).

In these subalpine meadow communities, the timing and the character of spring is influenced by two important factors: the warmth of the weather as the snowpack is melting, and the depth to which that snowpack accumulated during the winter. Because these two components are determined by weather conditions that prevail at different times of the year, there is little reason to expect them to covary in any strong and consistent way. Inconsistent variation should generate phenological mismatches, and natural selection should in turn produce organisms with some resilience mechanisms. For example, bees could store sufficient reserves in their fat bodies to ensure their survival if they emerge before flowers are available. Alternatively, they might be able to revert temporarily to the dormant state in which they overwintered. Flowers that have opened before pollinators are available, or during poor weather, may be able to extend their receptive lifespans (Fung & Thomson, 2017), and perennial plants that fail to be pollinated in one year may be able to present more flowers in the following year because they have been spared the cost of fruiting.

Although the previously reported trend of deteriorating pollination service across years (Thomson, 2010) has disappeared, the results of pollination supplementation experiments are hardly random. In particular, the recent data strongly reinforce the previously reported tendency for pollination deficits to decline during a flowering season (Fig. 2); indeed, for each year from 2015 to 2019, the *early* subset showed a significant deficit while the *late* subset did not. Unlike the pattern of deficits across years, which must be viewed as specific to the study species and locale, the within-season heterogeneity in pollination deficits may well be broadly applicable to plant–pollinator relationships, especially in seasonal environments. Indeed, direct measures of animal visitation rate in nearby subalpine meadows (Thomson, 1982) showed that visitation to 8/17 species, including *E. grandiflorum*, started out low and then increased during a species' period of bloom.

Of the hundreds of studies published on pollination limitation, most report a single value for a whole flowering season, leading implicitly to the interpretation that all members of the study population experienced some uniform level of pollination deficit. In cases such as *E. grandiflorum*, however, early- and late-flowering individuals experience fundamentally different pollination environments, and therefore would be expected to face different selection regimes for reproductive traits. This within-season heterogeneity is of small consequence to this study's question of whether pollination service is deteriorating, but for microevolutionary studies of selection, it may be an important consideration that is frequently overlooked.

Acknowledgements

Barbara Thomson assisted with all aspects of the work; she and David Andrews consulted on statistics. Jessica Forrest, David Inouye, Steve Johnson and anonymous reviewers provided valuable discussion, as did Jane Ogilvie, who also provided field assistance. Numerous assistants dissected fruits, including Maria Malzone, Alison Leslie, Matt Somers, Sadia Fazia, Jessica Zung, Hannah Fung and Dante Ravenhearst. This study was funded by US NSF (before 2000) and Canadian NSERC (after 2000).

ORCID

James D. Thomson https://orcid.org/0000-0002-1171-3589

References

Ågren J. 1996. Population size, pollinator limitation, and seed set in the self-incompatible herb *Lythrum salicaria*. *Ecology* 77: 1779–1790.

Aizen MA, Harder LD. 2007. Expanding the limits of the pollen-limitation concept: effects of pollen quantity and quality. *Ecology* **88**: 271–281.

Biesmeijer JC, Roberts SP, Reemer M, Ohlemüller R, Edwards M, Peeters T, Schaffers AP, Potts SG, Kluekers R, Thomas CD *et al.* 2006. Parallel declines in pollinators and insect pollinated plants in Britain and the Netherlands. *Science* **313**: 351–354.

Bos MB, Veddeler D, Bogdanski AK, Klein AM, Tscharntke T, Steffan-Dewenter I, Tylianakis JM. 2007. Caveats to quantifying ecosystem services: fruit abortion blurs benefits from crop pollination. *Ecological Applications* **17**: 1841–1849.

Burd M. 1994. Bateman's principle and plant reproduction: the role of pollen limitation in fruit and seed set. *Botanical Review* **60**: 83–139.

Forrest JRK, Thomson JD. 2011. An examination of synchrony between insect emergence and flowering in Rocky Mountain meadows. *Ecological Monographs* **81**: 469–491.

Fung HF, Thomson JD. 2017. Does lack of pollination extend flower life? *Journal of Pollination Ecology* **21**: 86–91.

García-Camacho R, Totland Ø. 2009. Pollen limitation in the alpine: a meta-analysis. *Arctic Antarctic and Alpine Research* **41**: 103–111.

Hegland SJ, Nielsen A, Lazaro A, Bjerknes A-L, Totland Ø. 2009. How does climate warming affect plant–pollinator interactions? *Ecology Letters* **12**: 184–195.

Inouye DW. 2000. The ecological and evolutionary significance of frost in the context of climate change. *Ecology Letters* **3**: 457–463.

Inouye DW. 2008. Effects of climate change on phenology, frost damage, and floral abundance of montane wildflowers. *Ecology* **89**: 353–362.

Knight TM, Steets JA, Ashman TL. 2006. A quantitative synthesis of pollen supplementation experiments highlights the contribution of resource reallocation to estimates of pollen limitation. *American Journal of Botany* **93**: 271–277.

Knight TM, Steets JA, Vamosa JC, Mazer SJ, Burd M, Campbell DR, Dudash MR, Johnston MO, Mitchell RJ, Ashman T-L. 2005. Pollen limitation of plant reproduction: pattern and process. *Annual Reviews of Ecology Evolution and Systematics* **36**: 467–497.

Larson BMH, Barrett SCH. 2000. A comparison of pollen limitation in flowering plants. *Biological Journal of the Linnean Society* **69**: 503–520.

Martin C. 2015. A re-examination of the pollinator crisis. *Current Biology* **25**: R811–R815.

Memmott J, Craze PG, Waser NM, Price MV. 2007. Global warming and the disruption of plant–pollinator interactions. *Ecology Letters* **10**: 710–717.

Miller-Rushing A, Inouye DW. 2009. Variation in the impact of climate change on flowering phenology and abundance: an examination of two pairs of closely related wildflower species. *American Journal of Botany* **96**: 1821–1829.

National Research Council of the National Academies. 2006. *Status of pollinators in North America.* Washington, DC, USA: National Academy Press.

Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, Kunin WE. 2010. Global pollinator declines: trends, impacts and drivers. *Trends in Ecology and Evolution* **25**: 345–353.

Rosenheim JA, Williams NM, Schreiber SJ. 2014. Parental optimism versus parental pessimism in plants: how common should we expect pollen limitation to be? *American Naturalist* **184**: 75–90.

Thomson JD. 1982. Patterns of visitation by animal pollinators. *Oikos* **39**: 241–250.

Thomson JD. 2001. Connecting pollination deficits and pollinator declines: can theory guide us? *Conservation Ecology* **5**: 6. [WWW document] URL http://www.consecol.org/vol5/iss1/art6.

Thomson JD. 2010. Flowering phenology, fruiting success, and progressive deterioration of pollination in an early-flowering geophyte. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **365**: 3187–3199.

Williams PH, Osborne JL. 2009. Bumblebee vulnerability and conservation worldwide. *Apidologie* **40**: 367–387.

Wilson P, Thomson JD, Stanton ML, Rigney LP. 1994. Beyond floral Batemania: sexual selection on pollination success. *American Naturalist* **143**: 283–296.

Yamagishi H, Allison TD, Ohara M. 2005. Effect of snowmelt timing on the genetic structure of an *Erythronium grandiflorum* population in an alpine environment. *Ecological Research* **20**: 199–204.