# A meta-analysis of small-plot trial data to examine the relationship between crop development and environmental conditions in canola

## CARP 2022.16 – Final Report

Principal Investigator:
Christiane Catellier, M.Sc., P.Ag.
Research Associate – Indian Head Agricultural Research Foundation (IHARF)
Correspondence: ccatellier@iharf.ca

Prepared for:
Saskatchewan Canola Development Commission
Western Grains Research Foundation

August 31, 2023

# Abstract

Canola emergence rates vary widely with management practices and field conditions, thus it is difficult to know the specific seeding density required to achieve the optimum plant population for maximizing canola yields. It would be beneficial for growers to know the precise range of emergence rates that can be expected based on their practices as well as the local environmental conditions. Meta-analysis is a powerful analytical tool where many independent data sets can be combined into a single analysis to provide a more accurate, wide-ranging interpretation of a research topic. The greatest benefit of meta-analysis for agronomic research is the greater statistical power that results from increased replication across a greater diversity of environments. The objective of this project was to utilize archived small-plot canola agronomic trial data and corresponding regional weather data to conduct a meta-analysis to examine the relationship between environmental conditions and canola emergence. The combined data set comprised agronomic and environmental data from 12 different projects conducted across a total of 47 site-years in Saskatchewan and Manitoba from 2013-2022. Single-variable regression and two-variable interaction models with mixed effects were used to examine the effects of individual management and environment variables and their interactions on the percent emergence of canola. The overall average percent emergence was 60.7%. The meta-analysis confirmed that the field emergence values frequently observed in canola production in western Canada, in the range of as low as 20-30% to as high as 80-90%, can be explained by measurable management and environmental variables. Seeding density, seeding date, seed-placed fertilizer, and average pre- and post-seeding air temperature all had negative effects on the percent emergence of canola, while pre- and post-seeding precipitation had positive effects on emergence. Significant interactions between non-correlated independent variables indicated that seeding date and average air temperature before and after seeding were the most influential variables, but the effects appeared to be likely related to soil moisture. Interacting effects between management and environmental variables were more likely under more ideal conditions of the most influential variables, specifically earlier seeding dates, lower average temperatures, and higher precipitation. Growers should be able to utilize the results of the meta-analysis to adjust seeding densities under certain conditions to achieve the optimum plant populations.

# Introduction

The optimum plant population recommended for maximizing canola yield has been precisely determined in recent years (Angadi et al. 2003, CCC 2021, Gan et al. 2016, Hartman & Jeffrey 2020, Shirtliffe & Hartman 2009). However, the specific seeding density required to achieve the optimum plant population is less certain, as emergence rates are highly variable and are influenced both by management practices and environmental conditions (Clayton et al 2004, Harker et al 2012, Hwang et al 2015). The effect of various management practices on canola emergence has been studied extensively, yet the effect of environmental variables has more often been discussed in the context of management, and rarely quantified on their own. It would be beneficial for producers to know the precise range of emergence rates that can be expected based on their practices as well as the local environmental conditions.

Meta-analysis is a powerful analytical tool where many independent data sets can be combined into a single analysis to provide a more accurate, wide-ranging interpretation of a research topic of interest. The greatest benefit of meta-analysis for agronomic research is the greater statistical power that results from increased replication across a greater diversity of environments (Philibert et al 2012, Eagle et al 2017). A recent example that is relevant to canola production is provided by Hartman and Jeffrey (2020), who conducted a meta-analysis to investigate the relationship between plant density and economically optimal yield in canola. There are many more examples of meta-analyses in agronomic research, and the analytical methods used are numerous and diverse.

The challenge in appropriately combining data from several studies for a meta-analysis is often a lack of standardization in experimental protocols and insufficient metadata reporting (Eagle et al 2017). However, in western Canada, there is an abundance of canola agronomic research being conducted in various locations in any year. The research topics are wide ranging, but experimental protocols, data collection, and crop management are fairly standard and consistent, especially among small-plot trials. Further, there is a large degree of collaboration among the applied research organizations conducting small-plot research, thus, simplifying the aggregation of trial data including necessary metadata. The inferential potential of a combined data set is enormous, as trials are conducted over several years and in many locations, covering a wide range of environmental conditions. Combined with local environmental data, a meta-analysis could be particularly insightful in evaluating the variability in canola emergence across the canola-growing region.

Thus, the objective of this project is to utilize archived small-plot canola agronomic trial data from across the production region and corresponding regional weather data to conduct a meta-analysis to examine the relationship between environmental conditions and canola emergence.

# Methods

## *Data collection*

Archived small-plot canola trial data was aggregated in collaboration with Agri-ARM organizations in Saskatchewan. Only publicly funded trial data was included in the data set. The trials were conducted in multiple locations in Saskatchewan and Manitoba in 2013-2022, and included data from 12 different projects/tests and total of 47 site-years. The trials included in the data set are described in Table A-1 (appendix). Unfortunately, several site-years of certain tests could not be included due to insufficient meta-data reporting.

Trials that contained the following obligatory plot-level data were included: 1) Seeding density (seeds per area), OR seeding rate (weight per area) and seed size (thousand seed weight); 2) Spring plant density assessment. In some cases, only a target seeding density was available, which was considered sufficient. Germination rate of the seed lot was not included in the seeding density calculation when it was not available but was assumed to be high (>95%) for all hybrid canola seed lots. Percent emergence was calculated for each plot using the seeding density and measured spring plant density.

Additional agronomic data was also collected and included: 1) Cultivar; 2) Previous crop; 3) Seeding date; 4) Row spacing; 5) Rates of seed-placed N, P, and S fertilizer; and 6) Treatment number and description, by trial. Also, maturity, fall stubble density, and yield data were included as additional response variables to explore if available. These variables may be examined in subsequent analyses but were not included this report.

Daily weather data (mean temperature and precipitation) was obtained from the nearest Environment Canada weather station for each trial site and year. Several new variables were calculated and explored in the analysis but ultimately, only select weather variables were retained that were most suitable and representative of environmental conditions affecting canola emergence: 1) Pre-seed temperature (average temperature 0-21 days before seeding), 2) Pre-seed precipitation (total precipitation from January 1 to seeding date), 3) Post-seed temperature (average temperature 0-21 days after seeding), and 4) Post-seed precipitation (total precipitation 0-14 days after seeding).

## *Statistical Analysis*

Single-variable and multiple regression with mixed effects modeling was used to examine the effect of individual management and environment variables and their interactions on the percent emergence of canola. Mixed effects models are an appropriate analytical tool for metanalysis as they compartmentalize and account for unbalanced replication across years, locations, and other influential independent variables. This is done through the specification of random effects. To understand the correct specification of random effects, we examine the structure of the data set. Experiments, or tests, are replicated at many sites over several years. Every combination of site and year encompasses different environmental conditions and is referred to as a site-year. A single replicate of a test within a site-year is a trial. There are several replicates of each treatment within a trial. Thus, the random effects are: 1) replicate within test within site-year, site-year, treatment within test, and test. The random effects were the same for every model.

The effects of some independent variables are not easily isolated if they are not distributed evenly across the random effects structure. The following variables were very unevenly replicated across sites or years and so were not included in the analysis: row spacing, seed size, cultivar (variety), and previous crop.

Data were analyzed with the R statistical program, version 4.2.2 (R Core Team 2022), using the *lme4* package (Bates et al. 2015) for fitting mixed-effects models, and the *lmerTest* package (Kuznetsova et al. 2017) for assessing model fit and treatment effects. First, single variable models were fitted for each independent variable to assess their effect on percent emergence of canola. Two-variable interaction models were then fitted for different combinations of non-correlated independent variables. Models with significant interactions were selected for further interpretation.

## Results and Discussion

### *Single variable models*

The overall average percent emergence was 60.7%, based on an intercept-only model with the same random effects structure as all other models. Nearly all variables had a significant effect on canola emergence individually (Table 1). The effects of seeding density, seeding date, seed-placed N, seed-placed P, and seed-placed S (Figure 1), and pre-seed temperature, post-seed temperature, pre-seed precipitation, and post-seed precipitation (Figure 2) on percent emergence of canola are shown along with all data points. Regression intercepts and co-efficients (Table 1) can be used to quantify the strength of the relationship and calculate the expected change in emergence with an increase in each of the variables.

The effects of seeding density and seed-placed fertilizer on canola emergence were relatively mild and consistent with previous findings. Percent emergence varied from approximately 50% to 70% over the range of seeding densities, and approximately 47% to 62% over the range of seed-placed fertilizer rates observed in the data set. The effects of pre- and post-seed precipitation were also consistent with expectations, where percent emergence increased with greater levels of precipitation both before and after seeding. Pre-seed precipitation was more influential than post-seed precipitation, varying from approximately 53% to 72% emergence over the range of values. Canola emergence varied from 57% to 73% over the range of post-seed precipitation observed in the data set.

Seeding date and temperature, meanwhile, had surprisingly large effects on canola emergence that were not exactly as expected. Canola emergence decreased from 80% at the earliest seeding date to 30% at the latest seeding date, and varied even more, from 90% to 20% emergence over the range of observed pre- and post-seeding average temperatures. Further, the quadratic effect was not statistically significant for these variables (not shown). Canola emergence is expected to be positively affected by warmer average temperatures, however it is likely that the lowest average temperatures observed in this data set were adequate for optimum canola emergence. Similarly, we would expect canola emergence to increase with later seeding dates due to potentially warmer temperatures. As both later seeding dates and increasingly higher average temperatures were detrimental to canola emergence, we can infer that the effect is potentially related to soil moisture.

Table 1. Description of each independent variable and results of the tests of significance for the single variable models assessing the effect on canola percent emergence.

| Independent variable | Range | Mean | Regression Intercept | Regression co-efficient | Pr(>\|t\|) |
|---|---|---|---|---|---|
| Seeding density (seeds m$^{-2}$) | 30 – 200 | 119 | 75.1 | -0.124 | <0.001 |
| Seeding date (Julian) | 124 – 161 | 137 | 251 | -1.378 | <0.001 |
| Seed-placed N (kg ha$^{-1}$) | 0 – 30 | 4.46 | 62.6 | -0.513 | <0.001 |
| Seed-placed P (kg ha$^{-1}$) | 0 – 100 | 15.9 | 62.5 | -0.126 | 0.002 |
| Seed-placed S (kg ha$^{-1}$) | 0 – 25.6 | 1.91 | 61.3 | -0.450 | 0.003 |
| Pre-seed temperature (°C) | 3.73 – 16.1 | 9.14 | 110 | -5.16 | <0.001 |
| Pre-seed precipitation (mm) | 8.10 – 195 | 69.3 | 53.3 | 0.095 | 0.017 |
| Post-seed temperature (°C) | 9.98 – 17.1 | 13.5 | 141 | -5.99 | <0.001 |
| Post-seed precipitation (mm) | 0 – 111 | 24.6 | 57.3 | 0.139 | 0.152 |

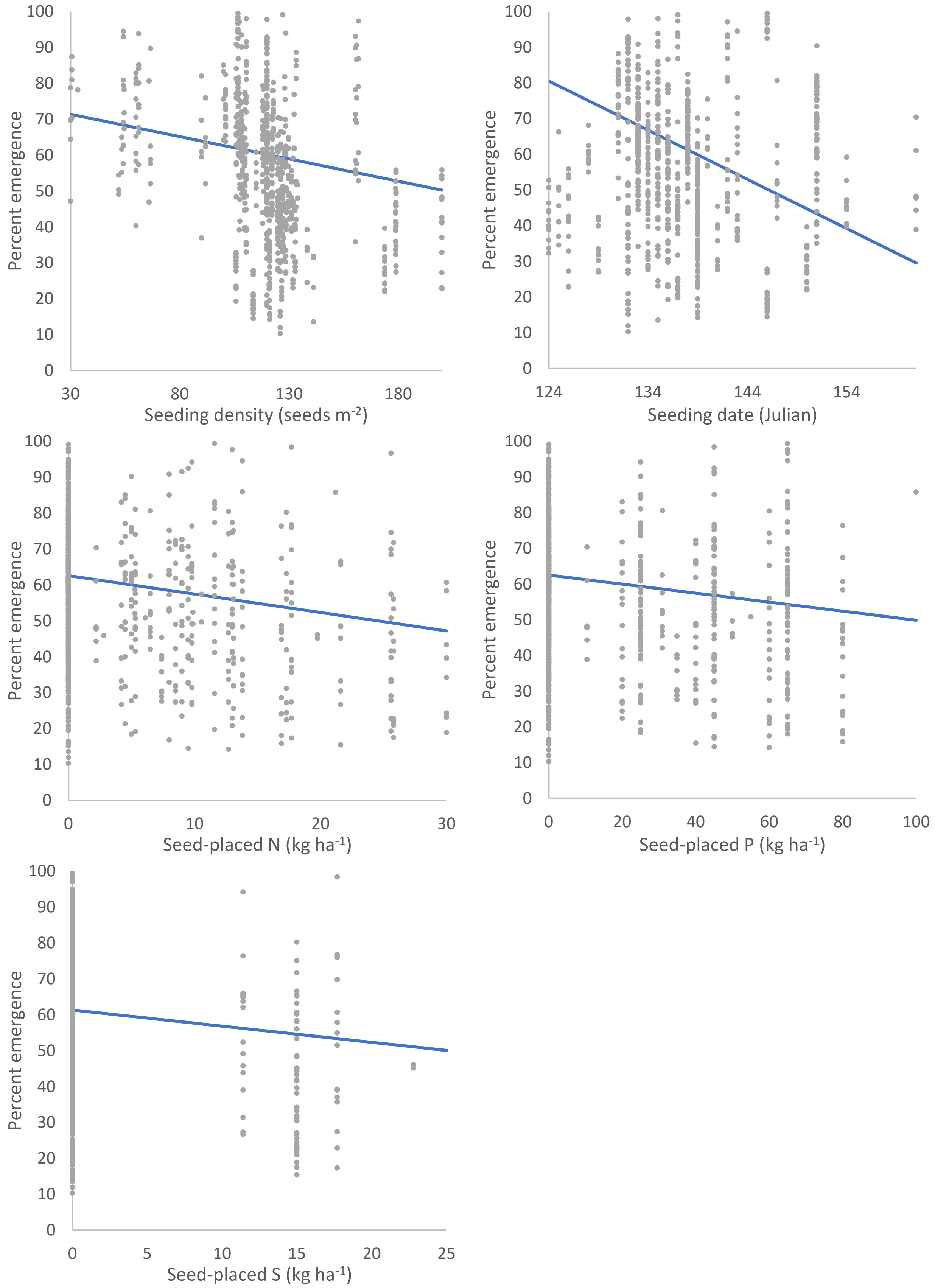

Figure 1. The individual effect of management variables on canola percent emergence. All data points are shown along with the regression line. Regression co-efficients and tests of significance are shown in Table 1.

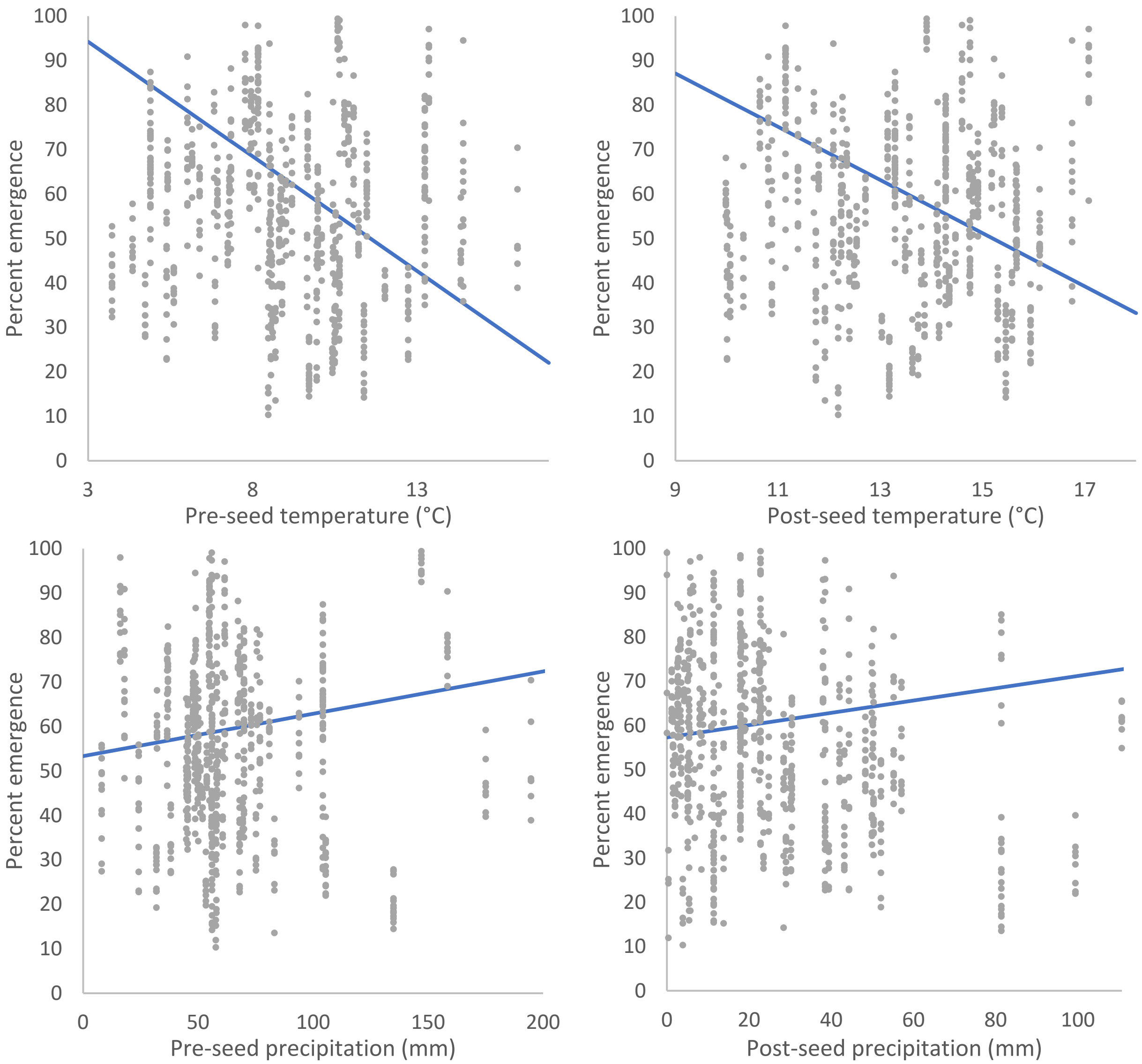

Figure 2. The individual effect of environmental variables on canola percent emergence. All data points are shown along with the regression line. Regression co-efficients and tests of significance are shown in Table 1.

### *Correlation among independent variables*

It is important to note correlations between the independent variables when interpreting single variable models and when selecting interaction models. A correlation graph of all independent variables is shown in Figure 3, though it should be noted that some level of correlation is accounted for in the random effects. When interpreting single variable models, it is difficult to separate the effects of correlated individual variables, they are likely confounded. For example, the effects of seed-placed N, seed-placed P, and seed-placed S on percent emergence are all similar (Figure 1), but we see that the three variables are also highly correlated (Figure 3). In this example, we are unable to differentiate between the effects of the three variables. We can also imagine how there may be some correlated/confounding variables that were not measured or included in the analysis, but the assumption is that these are mostly accounted for by the specification of random effects. In regards to the two-variable interaction models, correlated independent variables should not be included in the same model as it is impossible to separate their individual variances and so the coefficients and tests of significance are not reliable.

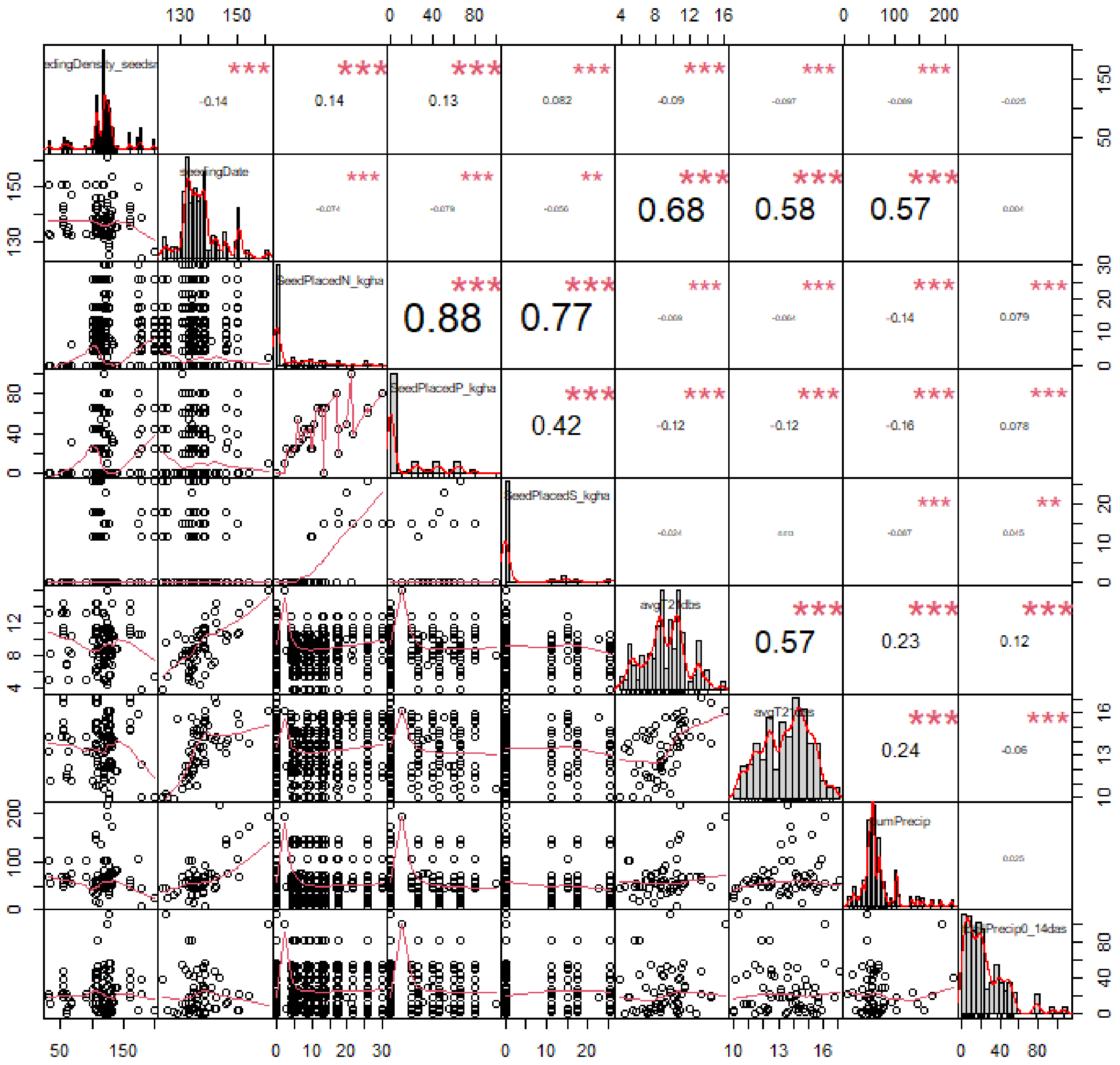

Figure 3. Correlation graph of all independent variables included in the analysis. Values in top left side of graph indicate correlation coefficients prior to  with random effects and so the significance values indicated by the red stars are not relevant.

## *Two-variable interaction models*

Two-variable interaction models were then fitted for combinations of non-correlated independent variables. Models that included seeding date with any environmental variable, two fertilizer variables, or two environmental variables were not fitted to avoid issues with correlated independent variables. All other combinations of variables were fitted and models with significant interactions between the two variables were selected for further interpretation. Regression model co-efficients and tests of significance for the models with significant interactions are shown in Table 2.

Seed-placed N, seed-placed P, and seed-placed S all had significant interactions with seeding date (Figure 4). The three fertility variables are correlated and had nearly identical relationships with seeding date so the effects of each could not be differentiated. Consistent with the single-variable model, canola emergence was lower overall when seeded at a later date compared to seeded early. The interaction between seeding date and seed-placed fertilizer showed that seed-placed fertilizer had little effect on canola emergence when seeded later. When seeded earlier, there was a significant decrease in emergence with increasing rates of seed-placed fertilizer. The later seeding date appears to have a stronger effect on canola emergence than seed-placed fertilizer. The emergence rate is already significantly lower with later seeding dates and so seed-placed fertilizer simply has a relatively lower influence on emergence overall when seeding is late.

Seed-placed N, seed-placed P, and seed-placed S also all had significant interactions with pre-seed precipitation (Figure 5). Again, the effects of the three fertility variables were similar and could not be differentiated. As in the single-variable model, canola emergence was higher overall with greater pre-seed precipitation. The interaction between pre-seed precipitation and seed-placed fertilizer showed that seed-placed fertilizer had little effect on emergence when pre-seed precipitation was high. With lower amounts of pre-seed precipitation, emergence decreased significantly with increasing amounts of seed-placed fertilizer. This interaction is as expected, where we would expect less damage from seed-placed fertilizer with higher levels of soil moisture.

Seed-placed N, but not seed-placed P and seed-placed S, had a significant interaction with average post-seeding temperature (Figure 6). Similar to the seeding date interaction, emergence was significantly lower overall with higher post-seeding temperatures and so was relatively less affected by seed-placed N.

Seeding density also had significant interactions with seeding date, pre-seed temperature, and pre-seed precipitation (Figure 7). In general, canola percent emergence decreased with higher seeding density, consistent with the single-variable model. As with all other models, percent emergence was also higher with earlier seeding dates, lower average pre-seed temperatures, and higher total pre-seed precipitation. Under these more ideal conditions, canola emergence decreased significantly more with increased seeding density. Under less ideal conditions – later seeding date, higher pre-seed temperatures, and lower pre-seed precipitation – canola emergence was low overall and not affected by increased seeding density.

Table 2. Regression model co-efficients and tests of significance for selected two-variable interaction models with significant interactions between two non-correlated independent variables.

| Variable 1 | Variable 2 | Intercept | Coeffic (Var1) | Coeffic (Var2) | Coeffic (Interaction) | P(Var1) | P(Var2) | P (Interaction) |
|---|---|---|---|---|---|---|---|---|
| Seeding date | Seed-placed N | 264.5 | -1.46 | -3.85 | 0.025 | <0.001 | <0.001 | 0.001 |
| Seeding date | Seed-placed P | 261.5 | -1.44 | -1.02 | 0.007 | <0.001 | 0.003 | 0.008 |
| Seeding date | Seed-placed S | 256.4 | -1.41 | -3.11 | 0.020 | <0.001 | 0.005 | 0.014 |
| Seed-placed N | Pre-seed precip | 56.0 | -0.88 | 0.09 | 0.006 | <0.001 | 0.030 | <0.001 |
| Seed-placed P | Pre-seed precip | 55.8 | -0.21 | 0.09 | 0.001 | <0.001 | 0.028 | 0.006 |
| Seed-placed S | Pre-seed precip | 54.2 | -0.72 | 0.09 | 0.005 | <0.001 | 0.021 | <0.001 |
| Seed-placed N | Post-seed temp | 150.9 | -1.37 | -6.58 | 0.062 | <0.001 | <0.001 | 0.013 |
| Seeding density | Pre-seed precip | 49.0 | 0.02 | 0.38 | -0.002 | 0.715 | 0.002 | 0.013 |
| Seeding density | Pre-seed temp | 142.0 | -0.33 | -6.81 | 0.020 | <0.001 | <0.001 | <0.001 |
| Seeding density | Seeding date | 383.3 | -1.23 | -2.22 | 0.008 | <0.001 | <0.001 | <0.001 |

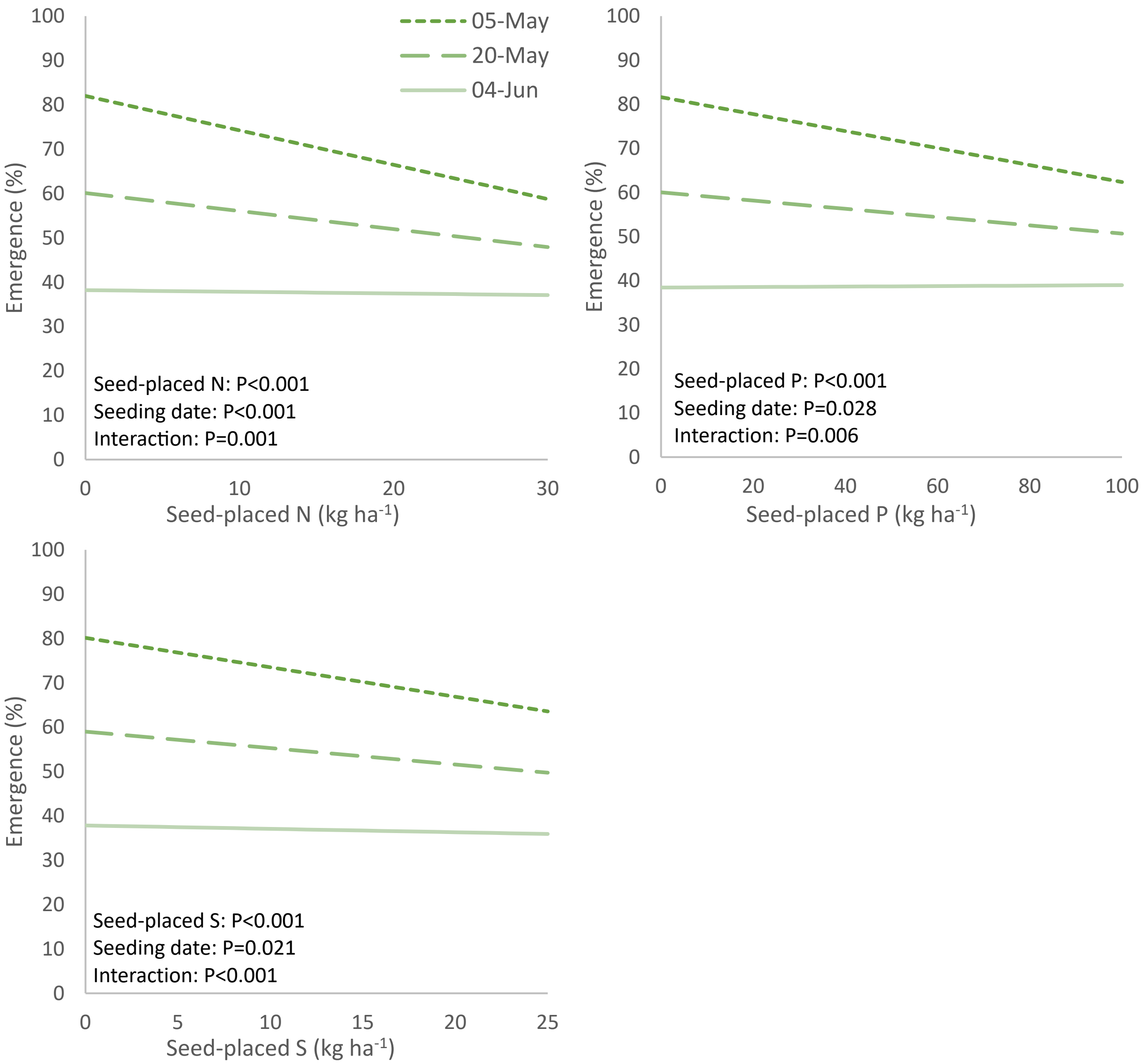

Figure 4. The interacting effect of seeding date with seed-placed fertilizer on canola percent emergence.

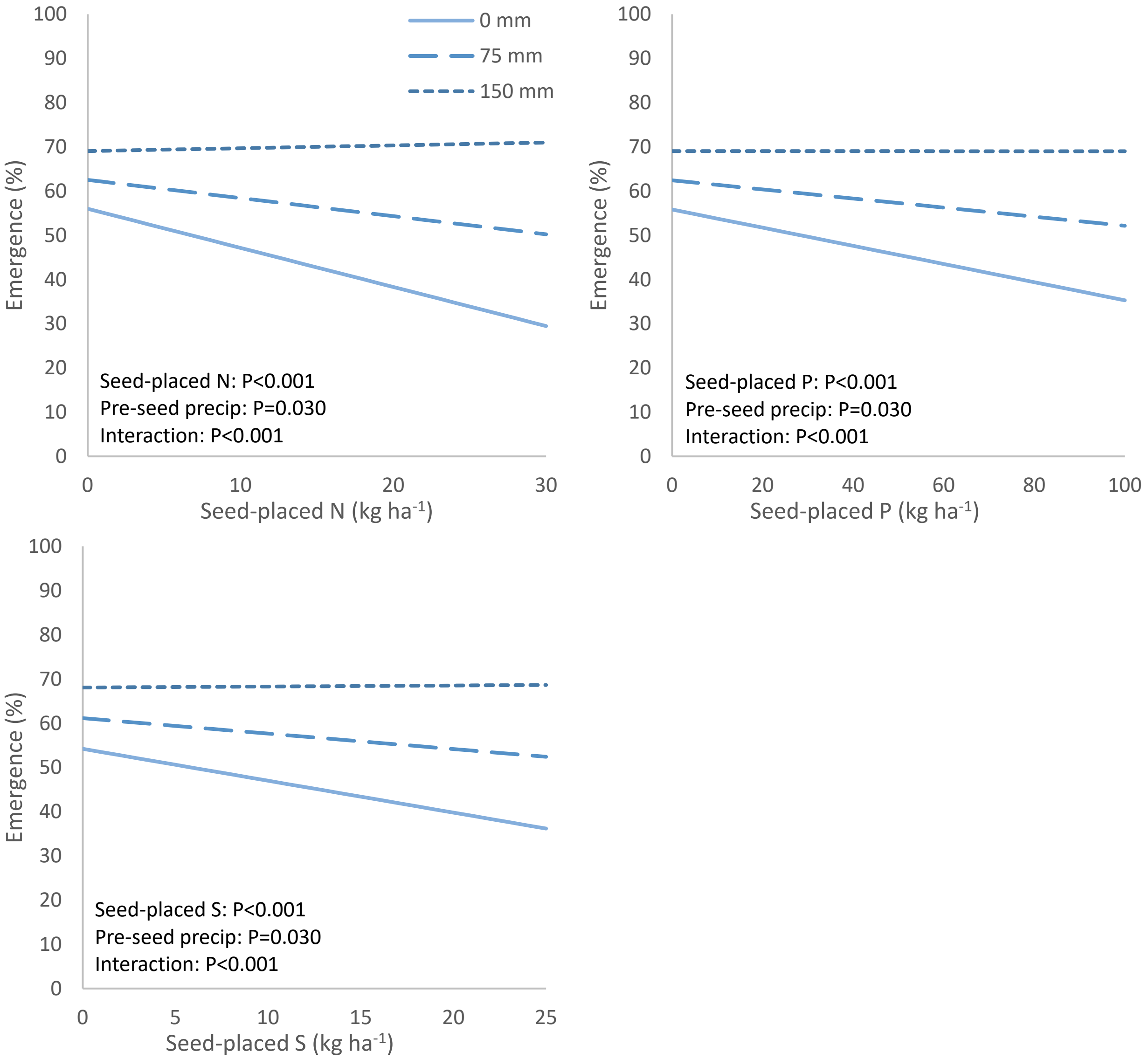

Figure 5. The interacting effect of pre-seed precipitation with seed-placed fertilizer on canola percent emergence.

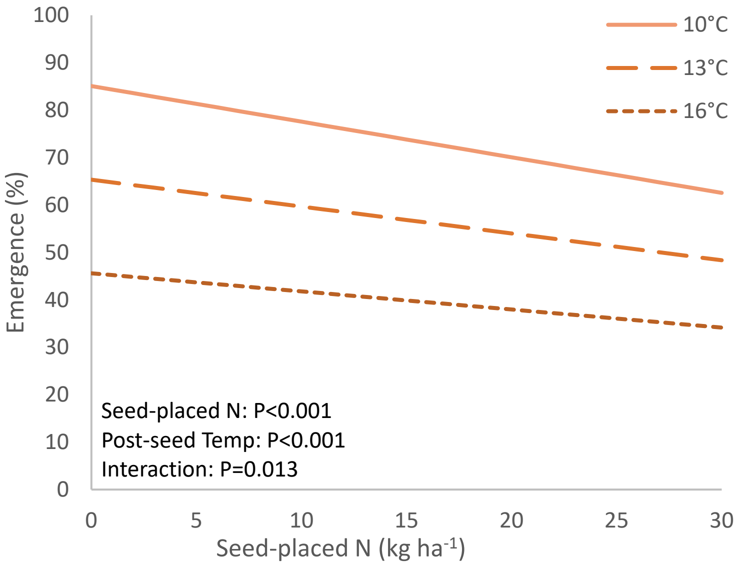

Figure 6. The interaction of seed-placed N with average post-seeding temperature on canola percent emergence.

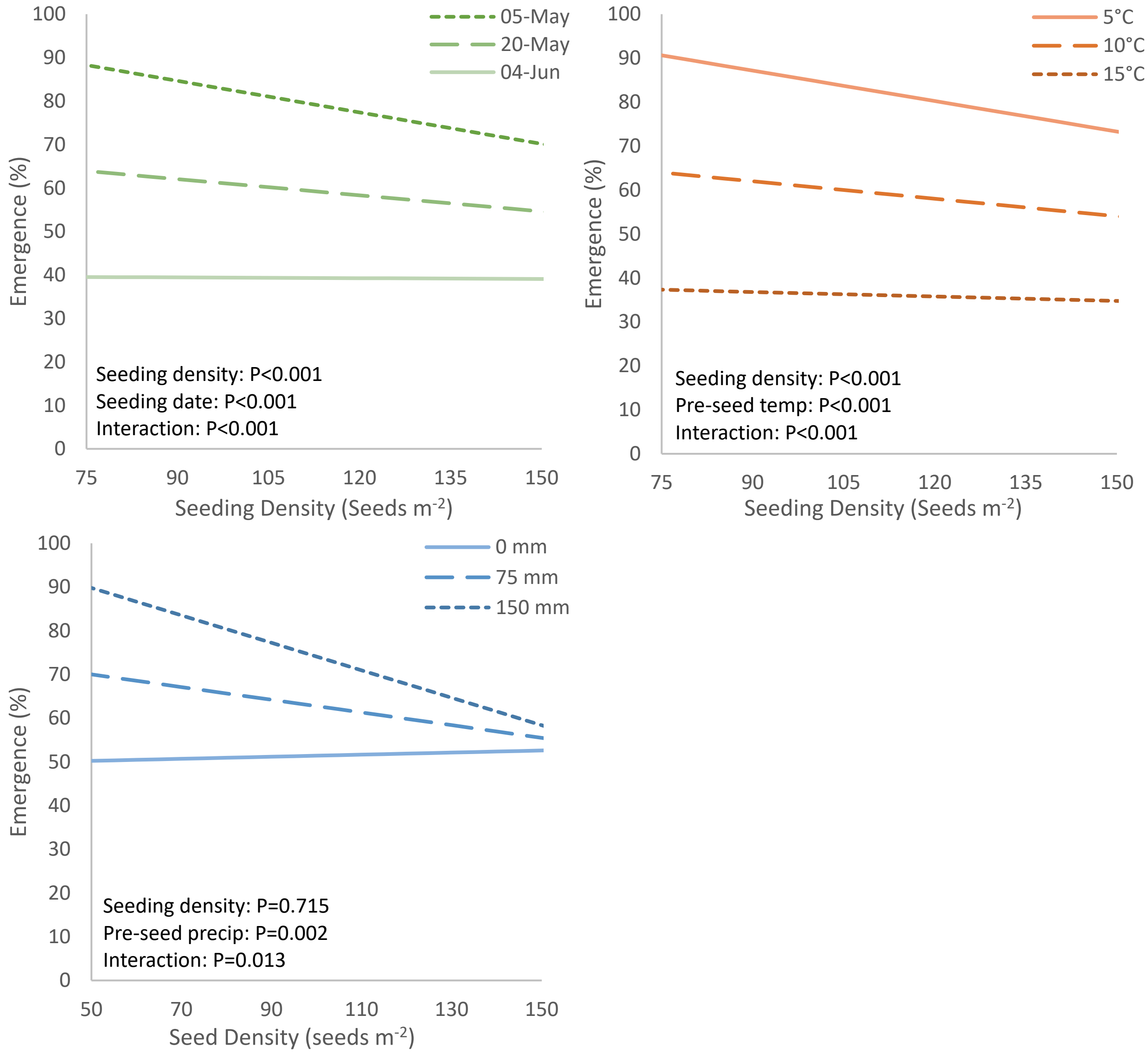

Figure 7. The interaction of seeding density with seeding date, pre-seed temperature, and pre-seed precipitation on canola percent emergence.

## Conclusions & Recommendations

The meta-analysis confirmed that the field emergence values frequently observed in canola production in western Canada, in the range of as low as 20-30% to as high as 80-90%, can be explained by measurable management and environmental variables. The large number of site-years included in the data set provide confidence that the data encompassed an accurate and representative range of spring field conditions. Seeding date and average air temperature before and after seeding were the most influential variables in the meta-analysis, but the effects appeared to be likely related to soil moisture, which was not examined in this project. Interacting effects between two variables were more likely under more ideal conditions of the most influential variables, specifically earlier seeding dates, lower average temperatures, and higher precipitation.

To confirm absolute emergence values that could be applied directly to commercial fields, similar data could be obtained from commercial fields. It would also be beneficial to include soil moisture and soil temperature as independent variables in any future studies if possible. As was discussed, the effects of air temperature on canola emergence appear to be related to soil moisture. Further, air temperature was used as a proxy for soil temperature yet the results suggest that air temperature and soil temperature could have contrasting effects on canola emergence.

An improvement in plant establishment has been identified as a key crop management area that will help achieve higher canola yield, profitability, and efficiency of inputs (CCC 2014). The strategic plan indicates that maximizing production will come from a more customized approach to agronomy, where targeted and relevant agronomic information and advice will be provided more specifically based on region, soil zone, and farm operation. Results of this study can be easily applied to commercial canola production. Though absolute percent canola emergence values may differ between small-plot trials and commercial fields, the relative reduction in emergence that is a result of the management and environmental variables is expected to be similar. Thus, to achieve the recommended plant densities to maximize canola yields, seeding rates may need to be increased when seeding at a later seeding date, when high average temperatures or low precipitation are observed before seeding, or expected after seeding. A better understanding of how environmental conditions affect canola emergence will allow producers to adjust their management based on the conditions within their operation. Producers will become more profitable by maximizing the value of their seed investment, and by more consistently achieving the target plant population recommended for optimizing productivity.

### *Extension and communication activities*

Extension and communication of the project results are forthcoming and will be conducted by the principal investigator in collaboration with canola industry extension and communication specialists.

### *Acknowledgements*

The principal investigator acknowledges the previous research efforts of project leads and all research personnel involved in collecting and managing the data included in the meta-analysis. Metadata reporting in particular was important for amalgamating data from various projects and sources. The final data set relied heavily on data sourced from IHARF, for projects led by Chris Holzapfel, due to detailed and accurate metadata reporting. Assistance from IHARF summer student Courtney Nell in compiling the full data set was also appreciated.

Funding for this project was provided by the Saskatchewan Canola Development Commission (SaskCanola) and the Western Grains Research Foundation (WGRF) under the Canola Council of Canada's (CCC) - Canola Agronomic Research Program (CARP).

## Literature Cited

Angadi et al. (2003). Yield adjustment by canola grown at different plant populations under semiarid conditions. Crop Sci 43:1358-1366.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67:1-48.

Canola Council of Canada (2014). 52 by 2025: How we'll get there. [Online] https://www.canolacouncil.org/about-us/strategic-plan/

Canola Council of Canada (2021). Canola Encyclopedia: Plant Populations. [Online] https://www.canolacouncil.org/canola-encyclopedia/plant-establishment/plant-populations/

Clayton et al. (2004). Fall and spring seeding date effects on herbicide-tolerant canola (Brassica napus L.) cultivars. Can J Plant Sci. 84:419-430.

Eagle et al. (2017). Meta-Analysis Constrained by Data: Recommendations to Improve Relevance of Nutrient Management Research. Agronomy Journal 109:2441-2449.

Gan et al. (2016). Canola seed yield and phenological responses to plant density. Can J Plant Sci 96:151-159.

Harker et al. (2012). Seeding depth and seeding speed effects on no-till canola emergence, maturity, yield and seed quality. Can. J. Plant Sci. 92: 795–802.

Hartman & Jeffrey (2020). Estimating the economic optimal target density of hybrid canola based on data from a western Canadian meta-analysis. Can J Plant Sci 101:393-407.

Hwang, et al. (2015). Effect of seeding date and depth, seed size and fungicide treatment on Fusarium and Pythium seedling blight of canola. Can J Plant Sci. 95:293-301.

Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). “lmerTest” Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software 82:1-26.

Philibert et al. (2012). Assessment of the quality of meta-analysis in agronomy. Agriculture, Ecosystems & Environment 148:72-82.

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Shirtliffe & Hartman (2009). Determining the economic plant density in canola [Online] https://www.saskcanola.com/research/determining-the-economic-plant-density-in-canola

# Appendix

Table A-1. Description of archived data sources utilized in the meta-analysis. Every site-year had four replicates of each treatment which were included individually in the data set.

| Test Name | No. treatments | Funder | Principal Investigator | Site-Years | No. Site-years | Total No. Observations |
|---|---|---|---|---|---|---|
| Effects of genetic sclerotinia tolerance and foliar fungicide applications on incidence and severity of sclerotinia stem rot infection in canola | 8 | CARP/SCDC | Chris Holzapfel | Brandon 2013-2015, Indian Head 2013-2015, Melfort 2013-2015, Melita 2013 & 2015, Outlook 2013-2015 | 14 | 448 |
| Investigating wider row spacing in no-till canola: implications for side-banded nitrogen fertilizer | 20 | CARP/SCDC | Chris Holzapfel | Indian Head 2013-2015 | 3 | 240 |
| Investigating wider row spacing in no-till canola: implications for seeding rate recommendations | 20 | CARP/SCDC | Chris Holzapfel | Indian Head 2013-2015 | 3 | 240 |
| Safe rates of side-banded and seed-placed phosphorus in canola | 11 | ADOPT | Chris Holzapfel | Indian Head 2015 | 1 | 44 |
| The impact of Lumiderm over standard canola seed treatments on flea beetle control and plant vigor | 10 | ADOPT | Mike Hall | Indian Head 2015, Melfort 2015, Scott 2015 | 3 | 120 |
| Enhancing canola production with improved phosphorus fertilizer management | 15 | CARP/SCDC | Stewart Brandt/ Jessica Pratchler | Indian Head 2016-2018, Melfort 2016-2018, Scott 2016-2018 | 9 | 540 |
| Demonstrating 4R phosporus principles in canola | 7 | ADOPT | Chris Holzapfel | Indian Head 2017 | 1 | 28 |
| Pre-harvest options for straight-combining canola | 10 | CARP | Chris Holzapfel | Indian Head 2017-2019, Melfort 2017-2019, Melita 2017-2019, Scott 2017-2019 | 12 | 480 |
| Seed-placed phosphorus fertilizer forms and *Penicillium bilaii* effects on canola emergence, P uptake and yield | 10 | ADOPT | Chris Holzapfel | Indian Head 2018 | 1 | 40 |
| Optimal seed size rate based on seed size in canola | 12 | SCDC | Christiane Catellier | Indian Head 2018, Melfort 2018, Outlook 2018, Scott 2018, Yorkton 2018 | 5 | 240 |
| Canola seed safety and yield response to novel P sources in Saskatchewan soils | 13 | ADOPT/Fertilizer Canada | Chris Holzapfel | Indian Head 2020-2022, Melfort 2021-2022, Outlook 2021, Redvers 2021, Scott 2020-2022, Swift Current 2020-2022, Yorkton 2020-2021 | 15 | 780 |
| Fall rye cover crop effects on canola establishment and response to nitrogen | 10 | ADOPT/Fertilizer Canada | Chris Holzapfel | Indian Head 2021-2022 | 2 | 80 |