Validation of predictive empirical weed emergence models of *Abutilon theophrasti* Medik based on intercontinental data

Valle Egea-Cabrero¹, Kevin Bradley², Isabel Calha³, Adam S. Davis⁴, Jose Dorado⁵, Frank Forcella⁶, John L Lindquist⁷, Christy L Sprague⁸, Jose L Gonzalez-Andujar^{1*}

Affiliations:

¹ Department of Crop Protection, Instituto de Agricultura Sostenible (CSIC), Córdoba, Spain.

² Department of Plant Science, University of Missouri, Columbia, MO, USA.

³ Instituto Nacional de Investigação Agrária e Veterinária, Oeiras, Portugal

⁴ Department of Crop Sciences, University of Illinois, Urbana, IL, USA.

⁵ Department of Plant Protection, Instituto de Ciencias Agrarias (CSIC). Madrid, Spain.

⁶ Department of Agronomy and Plant Genetics, University of Minnesota, St Paul, MN, USA

⁷ Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE, USA.

⁸ Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI, USA.

***Correspondence to:** andujar@cica.es

Short running title: Validation of predictive empirical weed emergence models

Summary

Emergence is possibly the most important demographic event for weed species, as it precedes and determines growth, survival and success. Good weed management relies on the proper timing of weed control practices in relation to weed emergence dynamics. Therefore, the development of models that predict the timing of emergence may help provide growers with tools to make better weed management decisions. The aim of this study was to validate and compare two previously published predictive empirical thermal time models of the emergence of *Abutilon theophrasti* growing in corn with data sets from the USA and Europe, and test the hypothesis that a robust and general weed emergence model can be developed for this species. Studies conducted in different and widely dispersed locations with the same species would be useful to assess whether or not weed species follow globally predictable patterns of emergence. Previously developed Weibull and Logistic models were validated against new data sets collected from 11 site-years, using four measures of validation. Our results indicated that predictions made with the Weibull model were more reliable than those made with the Logistic model but still contained appreciable biases to be considered a general applicable model of emergence for *A. theophrasti*. Our findings highlight the need to develop more accurate models if the ultimate goal is to make more precise predictions of weed seedling emergence globally to provide growers with tools to make better weed management decision.

Keywords: Weibull model, Logistic model, soil temperature, thermal model, velvetleaf

1. Introduction

Within the weed life-cycle, emergence is possibly the most important demographic event for an annual plant, determining its survival and success (Forcella *et al.*, 2000).

Moreover, predictive models of weed emergence can be useful tools to help farmers with timely attempts at chemical and physical weed control. Empirical models based on thermal and hydrothermal time have been developed to predict cumulative weed emergence of many species (Gonzalez-Andujar *et al.*, 2016). These models appear to be accurate enough for localized prediction purposes (e.g. Izquierdo *et al.*, 2009; Royo-Esnal *et al.*, 2010; Zambrano-Navea *et al.*, 2003).

Grundy *et al.* (2003) suggested the possibility of developing global models that allow their use in different locations. However, little information exists as to the feasibility of this suggestion. Existing evidence suggests the necessity of developing new site-specific predictive models (e.g. Dorado *et al.*, 2009a). Studies conducted in different and widely dispersed locations with the same species would be useful to assess whether or not weed species follow globally predictable patterns of emergence (Grundy, 2003).

Maize (*Zea mays* L.) is one of the world's staple crops. Its potential yield is reduced by several factors, with weeds being one of the most important. Among the major weed species that infest maize is *Abutilon theophrasti* Medik., which can reduce production by up to 80% (Lindquist *et al.*, 1996). Although *A. theophrasti* is native to eastern Asia, it has spread throughout North America (Warwick & Black, 1988) and Europe (Follak *et al.*, 2014) where it continues to be a major management concern for growers.

The aim of this study was to validate and compare two predictive empirical thermal time models (Weibull and Logistic) of *A. theophrasti* emergence with data sets from the USA and Europe, and test the hypothesis that a robust and general weed emergence model can be developed for this species. *Abutilon theophrasti* is especially appropriate for such an analysis as it responds to temperature and thermal time for seed germination and emergence in a highly uniform manner across populations from disparate geographic origins (Loddo *et al.*, 2013).

Material and Methods

Study sites

Validation was performed with data sets collected in the USA (Illinois, Michigan, Minnesota, Missouri, and Nebraska), Spain (Arganda) and Portugal (Golegã) (Table 1). Daily microclimate data were collected at each experimental site.

In the experiments conducted in the USA between 2012 and 2013, four 1-m² plots were located at each experimental site. Each plot contained four evenly spaced rows (25 cm between rows) with 30 seeds of *A. theophrasti* sown in each row. Seeds were sown in October or November each year. The following spring, seedling emergence was recorded non-destructively twice weekly until there were 2-4 plants per row. During this time, all newly emerged seedlings were marked with toothpicks to which a waterproof label with an identification number was attached. Once there were at least 2-4 seedlings per row, seedlings were thinned to leave the original 2-4 plants. Later emerging seedlings were counted, recorded and removed immediately.

Field experiments between 2005 and 2007 were conducted in Spain and Portugal. Monitoring of seedling emergence was performed in 42 quadrats (66 cm × 33 cm) in Arganda (Spain) and 28 quadrats in Golegã (Portugal). A complete description of methods is described by Dorado *et al.* (2009a).

Data from these experiments were applied to validate two thermal time (TT) emergence models developed for *A. theophrasti*. The first model used a Weibull function for cumulative emergence (E). It was developed by Dorado *et al.* (2009a) with data from Spain (Madrid).

$$E = 99.8(1 - \exp(-(0.006TT)^{1.3})) \quad (1)$$

The second model employed a Logistic function. It was developed by Myers *et al.* (2004) with data from the USA (Pennsylvania):

$$E=100/(1+\exp(18.86-3.21\ln(TT))) \quad (2)$$

To calculate TT, a base soil temperature of 9°C was used by both Dorado *et al.* (2009b) and Myers *et al.* (2004) for *A. theophrasti* seedling emergence. The former research team measured soil temperature at 3 cm depth and the latter at 2.5 cm. Any error associated with the slight difference in soil depth and temperature measurements was assumed to be negligible. The same base soil temperature and soil depth was used for the analysis reported here.

A base temperature of 9°C also was used by Hartzler *et al.* (1999) to calculate TT for *A. theophrasti* in Iowa, USA. Although seed germination of *A. theophrasti* can occur at minimum temperatures of 3 to 5°C (Loddo *et al.*, 2013; Mester & Buhler, 1991; Dorado *et al.*, 2009b), rapid germination does not occur until temperatures exceed 8°C (Mester & Buhler, 1991; Leon *et al.*, 2004).

Validation methods

One of the stages in the modeling process is the validation of the model. Diverse approaches and techniques have been employed in the literature for model validation (Tedeschi, 2006). Mainly, simple regression analysis is performed between mean model-outputs and real-system measurements as paired observations, and then the R^2 commonly is used to assess the relationship between observed and predict values (e.g. Myers *et al.*, 2004; Yousefi *et al.*, 2014). However, high values of R^2 (i.e., close to one) do not guarantee that the model fits the data well. Accordingly, here, combinations of four measures were used to validate models, the R^2 and three measures of deviation: Modeling

Efficiency (ME), Coefficient Error (CE) and Coefficient of Model Determination (CD) (Yang *et al.*, 2009).

ME is a dimensionless statistic and is similar to the Pearson correlation coefficient (Loague & Green, 1991).

$$ME = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (3)$$

where y_i is the i th observed value, $\hat{y}_i$ is the i th value predicted by the adjusted regression line, $\bar{y}$ is the arithmetic mean of the observed values, and z_i is the i th value predicted by the model to be validated. In a perfect fit ME would have a value equal to one. The lower boundary is negative infinity, and if ME is lower than zero, the model-predicted values are worse than the observed mean.

CE is a relative average measure of the absolute differences, which is expressed as a proportion of the mean of the observed values. A zero value for CE represents perfect agreement between predicted and observed values. It has been used to describe the adjustment of the model to the observed data and to compare the accuracy of the outputs of a model with different variables (Yang *et al.*, 2009).

$$CE = \frac{\sum|y_i - z_i|}{n/\bar{y}} \quad (4)$$

Where y_i , $\bar{y}$ and z_i are as indicated above, and n is the number of observations.

CD indicates the proportion of the total variation of the observed data explained by the predicted data. Like the ME, with a perfect fit CD would be one. A CD value close to one indicates an improvement in the predictions of the model. $CD > 1$ is an indicator of under prediction and if $CD < 1$ of over prediction (Tedeschi, 2006).

150
$$CD = \frac{\sum(y_i - \bar{y})}{\sum(z_i - \bar{y})^2} \quad (5)$$

151 Where y_i , z_i , and $\bar{y}$ are as noted above.

152

153 **Results**

154 Weibull and Logistic TT models were validated against data sets collected from 11 site-
155 years. R^2 values for the Weibull predictions ranged from 0.61 to 0.99 and averaged 0.80
156 across all localities. These values were higher than those arising from Logistic models,
157 which ranged from 0.30 to 0.97 and averaged 0.69 (Table 2). R^2 values for the Weibull
158 model results exceeded those of the Logistic model for seven of the eleven site-years (Fig.
159 1).

160

161 The average ME values for Weibull and Logistic models were 0.15 and -0.71,
162 respectively (Table 2). Values close to one indicate good models, whereas models with
163 negative values are not recommended for use. ME was negative in only one case for the
164 Weibull model results, but was negative in four cases for the Logistic model results. In
165 the remaining six site-years, the Weibull and Logistic models were split equally in terms
166 of which model's ME value was closer to unity.

167

168 The CD index gives an indication of over- or under-prediction. In general, both models
169 produced bias. The Logistic model produced over-prediction ($CD < 1$) in 9 site-years and
170 under-prediction in 2 site-years, while the Weibull model showed under-prediction in 5
171 site-years and over-prediction in 6 site-years (Table 2).

172

173 A zero value for CE represents perfect agreement between predicted and observed values.

174 The CE values obtained for the Weibull model ranged from excellent (0.02) to very high

(1.52) with an average of 0.34 (Table 2), which means that the average residual is 34% of the average measured value. The Logistic model presented a slightly worse average CE value (0.46), and ranged from excellent (0.08) to high (0.84) (Table 2). Nevertheless, average CE values showed that the Weibull model was closer to nil than the Logistic model (Table 2).

Discussion

Good weed management relies on the proper timing of weed control practices in relation to weed emergence dynamics (Grundy, 2003). Therefore, the development of models that predict the timing of emergence may help provide growers with tools to make better weed management decisions. This can be especially critical in regions where, and at times when, species of weeds are resistant to multiple classes of herbicides (Beckie *et al.*, 2019). To be reliably operational, the models need to be validated by comparing their predictions with observed data. Here, we compared two models developed and validated in two continents with entirely independent observational data for the emergence of *A. theophrasti*. The Logistic model was considered appropriate to predict the emergence of *A. theophrasti* in the USA (Myers *et al.*, 2004), while the Weibull model was adequate for this species in the Iberian Peninsula (Dorado *et al.*, 2009a). In both of these cases, however, model performance (validation) was assessed by the respective research teams only with the coefficient of determination (R^2). Here, we tested the global predictive capacity of both types of models using a combination of different assessment metrics. The previous study of Myers *et al.* (2004) with the Logistic model reported R^2 values ranging from 0.91 to 0.96 for *A. theophrasti* across three site-years in Pennsylvania (USA). Using that model to examine data from 8 site-years from across the Corn Belt in the USA, our results found a range for R^2 from 0.30 to 0.94. Moreover, the average ME

value for the Logistic model was negative, which indicated poor performance (Table 2). Note that in the USA, five of eight ME values were negative (Table 2), whereas Logistic model behaviour was better in the Iberian Peninsula (i.e., no negative ME values). In contrast, the Weibull model presented an average positive ME value, while only one site showed a negative value within the USA (Table 2).

Our results showed average CD values close to 1 (the perfect fit) in both models. The Logistic model presented mostly over-prediction (i.e. $CD < 1$) in nine of eleven sites-years, while the Weibull model showed a more even distribution with over- and under-prediction in five and six of eleven site-years, respectively (Table 2).

The CE values for the Weibull model were closer to zero (i.e., close to perfect agreement between predicted and observed values) than those of the Logistic model in seven of the eleven cases. Moreover, average CE for the Weibull models was 0.34 compared to 0.46 for the Logistic model (Table 2). Thus, both in terms of frequency and on average, the Weibull model fit data better than the Logistic model.

Taken overall, these results indicate the Weibull function is a more reliable model than the logistic model but, nevertheless, it still is a model that contains appreciable biases if the ultimate goal is to make accurate predictions of weed seedling emergence globally. The superiority of the Weibull model compared to the Logistic model is likely a consequence of its greater flexibility (Calvo Haro *et al.*, 1994); e.g., it does not require symmetry on either side of its midpoint predictions.

Possibly there are many sources of error and uncertainty associated with the datasets we used other than temperature and thermal time, and these may appreciably affect emergence characteristics of local weed populations. The moisture levels, aeration, crop management practices, genetic differences among populations, etc., may affect weed seedling emergence differentially across locations (Tozzi *et al.*, 2014; Loddo *et al.*, 2019). Understanding these effects will be critical for developing universally applicable models of emergence. There is a clear need to perform more global experiments that permit more rigorous testing of the viability of a general model of emergence of *A. theophrasti*. In this sense also, other modelling approximations may be more appropriate to establish a general model that explains the deviations observed in site-year environments. For example, Davis *et al.* (2013) and Clay *et al.* (2014) used nonlinear mixed effects modelling approach for the regional emergence data for giant ragweed and sunflower, and Cao *et al.* (2013) and Gonzalez-Andujar *et al.* (2016) used a nonparametric approach for modelling *Bromus diandrus* and *Phalaris paradoxa* seedling emergence. The predictive capacity of the models will have an important impact on decision making by growers, therefore the need for more accurate models is critical.

Acknowledgement: VE-C and JLG-A were supported by FEDER (European Regional Development Funds) and the Spanish Ministry of Economy and Competitiveness funds (project AGL2015-64130-R). No conflicts of interest have been declared.

References

BECKIE HJ, BUSI R, BAGAVATHIANNAN V, & MARTIN SL (2019) Herbicide resistance gene flow in weeds: Under-estimated and under-appreciated. *Agric Ecosyst Environ* 283, 106566.

- 249 CALVO HARO RM, GONZALEZ-ANDUJAR JL & PEREZ BOADA S (1994)
 250 *Manual de modelos no lineales en los ámbitos agronómico, ganadero y forestal.*
 251 Ministerio de Agricultura, Pesca y Alimentación, Instituto Nacional de
 252 Investigación y Tecnología Agraria y Alimentaria Press. Madrid. Spain.
- 253 CAO R, FRANCISCO-FERNANDEZ M, ANAND A, BASTIDA F & GONZALEZ-
 254 ANDUJAR JL (2013) Modeling *Bromus diandrus* seedling emergence using
 255 nonparametric estimation. *J Agric Biol Environ Stat.* 18, 64–86.
- 256 CLAY SA, DAVIS A, DILLE A, LINDQUIST J, RAMIREZ AHM, SPRAGUE C,
 257 REICKS G & FORCELLA F (2014) Common sunflower seedling emergence
 258 across the U.S. Midwest. *Weed Sci.* 62:63-70.
- 259 DAVIS AS, CLAY SA, CARDINA J DILLE A, FORCELLA F, LINDQUIST J &
 260 SPRAGUE C (2013) Seed burial physical environment explains departures from
 261 regional hydrothermal model of giant ragweed (*Ambrosia trifida*) seedling
 262 emergence in US Midwest. *Weed Sci* 61, 415–421.
- 263 DORADO J, SOUSA E, CALHA I, GONZALEZ-ANDUJAR J & FERNÁNDEZ-
 264 QUINTANILLA C (2009a) Predicting weed emergence in maize crops under two
 265 contrasting climatic conditions. *Weed Res* 49, 251–260.
- 266 DORADO J, FERNÁNDEZ-QUINTANILLA C & GRUNDY AC (2009b) Germination
 267 patterns in naturally chilled and nonchilled seeds of fierce thornapple (*Datura*
 268 *ferox*) and velvetleaf (*Abutilon theophrasti*). *Weed Sci* 57, 155–162 .
- 269 FOLLAK S, ALDRIAN U & SCHWARZ M (2014) Spread dynamics of *Abutilon*
 270 *theophrasti* in Central Europe. *Plant Prot Sci* 50, 157–163.
- 271 FORCELLA F, BENECH ARNOLD RL, SANCHEZ R & GHERSA CM (2000)
 272 Modeling seedling emergence. *F Crop Res* 67, 123–139.
- 273 GONZALEZ-ANDUJAR JL, FRANCISCO-FERNANDEZ M, CAO R *et al.* (2016) A

- 274 comparative study between nonlinear regression and nonparametric approaches
275 for modelling *Phalaris paradoxa* seedling emergence. *Weed Research* 56, 367-
276 376.
- 277 GONZALEZ-ANDUJAR JL, CHANTRE GR, MORVILLO C, BLANCO AM &
278 FORCELLA F (2016) Predicting field weed emergence with empirical models
279 and soft computing techniques. *Weed Res* 56, 415–423.
- 280 GRUNDY A (2003) Predicting weed emergence: a review of approaches and future
281 challenges. *Weed Res* 43,1–11.
- 282 GRUNDY AC, PETERS NCB, RASMUSSEN IA *et al.* (2003) Emergence of
283 *Chenopodium album* and *Stellaria media* of different origins under different
284 climatic conditions. *Weed Res* 43, 163–176.
- 285 HARTZLER R, BUHLER D & STOLTENBERG D (1999) Emergence characteristics
286 of four annual weed species. *Weed Sci* 47: 578–584 .
- 287 IZQUIERDO J, GONZALEZ-ANDUJAR JL, BASTIDA F *et al.* (2009) A thermal time
288 model to predict corn poppy (*Papaver rhoeas*) emergence in cereal fields. *Weed*
289 *Sci* 57: 660–664 (2009).
- 290 LEON RG, KNAPP AD & OWEN MDK (2004) Effect of temperature on the
291 germination of common waterhemp (*Amaranthus tuberculatus*), giant foxtail
292 (*Setaria faberi*), and velvetleaf (*Abutilon theophrasti*). *Weed Sci* 52, 67–73.
- 293 LINDQUIST J, MORTENSEN D, CLAY S *et al.* (1996) Stability of corn (*Zea mays*)
294 velvetleaf (*Abutilon theophrasti*) interference relationships. *Weed Sci* 44, 309–
295 313.
- 296 LOAGUE K & GREEN RE (1991) Statistical and graphical methods for evaluating
297 solute transport models: Overview and application. *J Contam Hydrol* 7, 51–73.
- 298 LODDO D, SOUSA E, MASIN R *et al.* (2013) Estimation and comparison of base

- temperatures for germination of european populations of velvetleaf (*Abutilon theophrasti*) and jimsonweed (*Datura stramonium*). *Weed Sci* **61**, 443–451.
- LODDO, BOZIC D, CALHA IM et al. (2019) Variability in seedling emergence for European and North American populations of *Abutilon theophrasti*, *Weed Research*, **59**, 15–27.
- MESTER T & BUHLER D (1991) Effects of soil-temperature, seed depth, and cyanazine on giant foxtail (*Setaria faberi*) and velvetleaf (*Abutilon theophrasti*) seedling development. *Weed Sci* **39**, 204–209.
- MYERS MW, CURRAN WS, VANGESSEL MJ et al. (2004) Predicting weed emergence for eight annual species in the northeastern United States. *Weed Sci* **52**, 913–919.
- ROYO-ESNAL A, TORRA J, CONESA JA, FORCELLA F & RECASENS J (2010) Modeling the emergence of three arable bedstraw (*Galium*) species. *Weed Sci* **58**, 10–15.
- TEDESCHI LO (2006) Assessment of the adequacy of mathematical models. *Agric Syst* **89**, 225–247.
- WARWICK S & BLACK L (1988) The biology of canadian weeds. 90. *Abutilon theophrasti*. *Can J Plant Sci* **68**, 1069–1085.
- YANG J, GREENWOOD D, ROWELL D, WADSWORTH G & BURNS I (2000) Statistical methods for evaluating a crop nitrogen simulation model, N-ABLE. *Agric Syst* **64**, 37–53.
- YOUSEFI AR, OVEISI M & GONZALEZ-ANDUJAR JL (2014) Prediction of annual weed seed emergence in garlic (*Allium sativum* L.) using soil thermal time. *Sci Hortic (Amsterdam)* **168**, 189–192.
- ZAMBRANO-NAVEA C, BASTIDA F & GONZALEZ-ANDUJAR JL (2013) A

324 hydrothermal seedling emergence model for *Conyza bonariensis*. *Weed Res* **53**,
325 213–220.

Review Copy

Figure 1. The best and worst model fits obtained with the Weibull model (a, b) and Logistic model (c, d). Closed circles, observed data; solid line, model fit. ME: Modelling Efficiency; CD: Coefficient of model Determination; CE: Coefficient of Error; R^2 : Coefficient of determination

Table 1. Geographical location, soil type, and climatic conditions for the experimental sites

Site	Location	Average temperature (°C) and precipitation (mm)		Soil Type
Minnesota (USA)	45°68'N, 95°80'W	5.8	673	clay loam
Illinois (USA)	40°02'N, 88°14'W	10.9	1045	silty clay loam
Nebraska (USA)	41°09'N, 96°24'W	10.8	735	silty clay loam
Michigan (USA)	42°42'N, 84°28'W	8.2	785	sandy clay loam
Missouri (USA)	38°53'N, 92°22'W	12.5	1083	clay loam
Arganda (Spain)	40°19'N, 3°19'W	13.5	461	sandy loam
Golegã (Portugal)	39°24'N, 8°29'W	16.7	719	sandy loam

Table 2. Model performance statistics for Weibull (W) and Logistic (L) models. ME: Modelling Efficiency; CD: Coefficient of model Determination; CE: Coefficient of Error; R²: Coefficient of determination

	ME		CD		CE		R ²	
	W	L	W	L	W	L	W	L
MN 2012	1.00	-4.59	0.38	0.10	0.32	0.84	0.76	0.34
MN 2013	0.75	-1.02	1.74	0.53	0.25	0.79	0.89	0.30
IL 2012	1.00	-0.98	1.23	0.18	0.02	0.43	0.99	0.65
MO 2012	0.00	0.69	0.42	7.53	0.43	0.28	0.61	0.70
MI 2012	0.78	0.11	1.14	0.60	0.19	0.48	0.89	0.84
MI 2013	0.12	-4.1	0.25	0.05	0.30	0.82	0.72	0.34
NE 2012	-4.58	0.67	0.08	0.38	1.52	0.24	0.78	0.94
NE 2013	0.87	-1.15	0.76	0.26	0.11	0.57	0.91	0.73
AG 2005	0.85	0.70	1.27	0.55	0.09	0.11	0.92	0.83
GL 2006	0.49	0.96	1.13	0.94	0.26	0.09	0.69	0.97
GL 2007	0.41	0.93	1.27	1.05	0.25	0.08	0.62	0.94
Mean	0.15	-0.71	0.88	1.11	0.34	0.43	0.80	0.69

In bold CD-index indicating overprediction. MN:Minnesota; IL:Illinois; MO:Missouri; MI:Michigan; NE:Nebraska; AG: Arganda; GL:Golegã