Journal of Agricultural Science (2014), **152**, 254–262. © Cambridge University Press 2013
doi:10.1017/S0021859612001098

CROPS AND SOILS RESEARCH PAPER

A comparative study between non-linear regression and artificial neural network approaches for modelling wild oat (*Avena fatua*) field emergence

G. R. CHANTRE[1], A. M. BLANCO[2], F. FORCELLA[3], R. C. VAN ACKER[4], M. R. SABBATINI[1] AND J. L. GONZALEZ-ANDUJAR[5]*

[1] *Departamento de Agronomía/CERZOS, Universidad Nacional del Sur/CONICET, Bahía Blanca, Buenos Aires, 8000, Argentina*
[2] *Planta Piloto de Ingeniería Química, Universidad Nacional del Sur/CONICET, Bahía Blanca, Buenos Aires, 8000, Argentina*
[3] *USDA-ARS North Central Soil Conservation Research Laboratory, Morris, MN 56267, USA*
[4] *Department of Plant Agriculture, University of Guelph, Guelph, ON, N1G 2W1, Canada*
[5] *Instituto de Agricultura Sostenible (CSIC), Aptdo. 4084, 14080 Córdoba, Spain*

(*Received 30 August 2012; revised 18 December 2012; accepted 20 December 2012; first published online 23 January 2013*)

SUMMARY

Non-linear regression (NLR) techniques are used widely to fit weed field emergence patterns to soil microclimatic indices using S-type functions. Artificial neural networks (ANNs) present interesting and alternative features for such modelling purposes. In the present work, a univariate hydrothermal-time based Weibull model and a bivariate (hydro-time and thermal-time) ANN were developed to study wild oat emergence under non-moisture restriction conditions using data from different locations worldwide. Results indicated a higher accuracy of the neural network in comparison with the NLR approach due to the improved descriptive capacity of thermal-time and the hydro-time as independent explanatory variables. The bivariate ANN model outperformed the conventional Weibull approach, in terms of RMSE of the test set, by 70·8%. These outcomes suggest the potential applicability of the proposed modelling approach in the design of weed management decision support systems.

INTRODUCTION

Significant efforts have been made by the weed research community to develop accurate weed emergence models. These models aim to predict the time and magnitude of weed emergence flushes to assist farmers in determining the best time to apply a control measure. However, weed emergence in the field relies heavily on soil conditions and weed biology which, on some occasions, is dominated by as yet unravelled seed dormancy processes hindering the task of accurate prediction (Forcella *et al*. 2000; Batlla & Benech-Arnold 2007).

Usually, researchers have adopted both mechanistic and empirical approaches to predict weed emergence. Mechanistic models that provide an intimate description of the basic eco-physiological processes underlying seedling emergence (Colbach *et al*. 2011; Colbach & Mézière 2013; Gardarin *et al*. 2010, 2012) are desirable from an explanatory point of view. However, as Grundy (2003) remarked, these models are more difficult to develop and, although built on a biological process-based framework, they do not have the simplicity and flexibility that would be required for practical decision support, which are offered by weather-based models. In addition, reductionist approaches for weed emergence prediction still remain largely empirical in nature (Grundy 2003).

Empirical models to predict weed seedling emergence in agronomic systems are usually built on variables related to soil temperature and soil water potential. Soil microclimate-derived indices such as hydrothermal time, thermal time or hydro-time are commonly used for model development. They assume

* To whom all correspondence should be addressed. Email: andujar@cica.es

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

that emergence rates depend on the amount by which soil temperature and soil water potential exceed a given threshold (Bradford 2002). These approaches tend to make emergence predictions independent of weather conditions (e.g. air temperature and precipitation) and soil characteristics (e.g. texture and soil density), thus enabling the development of more general and widely applicable models.

Parametric non-linear regression (NLR) models based on S-shaped curves (e.g. Weibull) using thermal and hydro-thermal indices as explanatory variables have been used widely for weed emergence prediction (Forcella 1998; Schutte *et al.* 2008; Izquierdo *et al.* 2009; Bullied *et al.* 2012). In many cases, NLR models have demonstrated adequate representation to the observed data; however, they present several major limitations. Specifically, NLR models are sometimes not flexible enough to capture complex features in the explanatory variable, such as abrupt 'jumps' or heavy 'tails' (Cao *et al.* 2011). Moreover, observed cumulative emergence values obtained from consecutive monitoring approaches are not statistically independent. However, this is not considered explicitly in the weed science literature, where fitting the model is the goal regardless of whether the data are independent or not (Onofri *et al.* 2010; Cao *et al.* 2011).

Alternative modelling approaches based on survival analysis (Onofri *et al.* 2010, 2011), genetic algorithms (Haj Seyed Hadi & González-Andújar 2009) and artificial neural networks (ANNs) (Chantre *et al.* 2012) have also been developed. In particular, ANNs are semi-parametric tools that provide a practical and flexible modelling framework known for their capacity to describe highly non-linear relationships among variables, thus showing a high potential applicability in ecological systems (Lek & Guégan 1999). ANNs are massively interconnected parallel processors that allow the storage of experimental knowledge (training data) in order to use it for a predictive purpose. Some advantages of ANNs are: (i) effective implementation of a wide variety of non-linear mappings, (ii) admittance of input/output multivariate mapping, (iii) no requirement for a given *a priori* shape fitting function and (iv) requirement of fewer adjustable parameters than conventional multivariate techniques (*sensu lato*). Although ANNs have been used intensively to solve highly complex non-linear mapping problems in agronomical systems (Alvarez 2009), their application for modelling weed emergence remains largely unexplored (Chantre *et al.* 2012).

Wild oat (*Avena fatua* L.) is a world-wide invasive species that competes strongly with small grain cereal crops, causing significant yield losses when left unmanaged (Cousens *et al.* 1991). Consequently, a thorough understanding of the timing and magnitude of wild oat seedling emergence in the field is important for its control in crop systems. Some NLR models have been developed to explain the emergence of this species (González-Andújar *et al.* 2001; Martinson *et al.* 2007; Page *et al.* 2006; Moschini *et al.* 2009). More recently, models based on ANNs have been proposed for wild oat emergence prediction as an alternative to NLR models, with successful results (Chantre *et al.* 2012).

The objective of the present work was to perform a comparative study between ANNs and NLR approaches to model wild oat emergence using data gathered from locations in the United States, Canada and Australia.

MATERIALS AND METHODS

Field experimental data

Wild oat emergence data for a number of years and locations in Canada, USA and Australia were gathered (Table 1). Emergence data from North Dakota and Minnesota (USA) were derived from Martinson *et al.* (2007). Associated weather data were obtained from the following websites: http://ndawn.ndsu.nodak.edu/daily-table-form.html and http://climate.umn.edu/HIDradius/radius.asp (211886 Crookston). Seedling information from Montana (USA) came from Mickelson & Grey (2006) and weather data from http://www.sarc.montana.edu/php/weather. For Manitoba (Canada), seedling emergence data were reported (A. M. Marginet, personal communication) and further discussed by Bullied *et al.* (2003). Weather data came from the nearby official weather stations of Brandon and Cypress River, Manitoba (Canada). Web-based weather records were downloaded from http://climate.weatheroffice.gc.ca/advanceSearch/searchHistoricData_e.html. Lastly, data from South Australia were taken from Chauhan *et al.* (2006) and http://www.bom.gov.au/climate/data/. Up to 0·20 of the data for some weather variables were missing at Roseworthy, South Australia. In some cases, minimum air temperature was recorded but not maximum air temperature (or vice versa). To rectify this situation, days with any missing values were eliminated from a revised data set and linear regressions devised for

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

Table 1. *Description of site locations*

Location	Years	Soil characteristics	Coordinates and elevation	Tillage system
Crookston, Minnesota, USA	2002/2003	Clay loam, 5% OM	38°08′N-97°00′W 273 m asl	No-tillage
Fargo, North Dakota, USA	2002/2003	Clay, 4·5% OM	46°52′N-96°47′W 274 m asl	No-tillage
Huntley, Montana, USA	2000/2001/2002	Silty clay, 2·3% OM	45°53′N-108° 18′W 921 m asl	Conv. Tillage
Brandon, Manitoba, Canada	2000	Clay loam, 5% OM	49°54′N-99°57′W 388 m asl	No-tillage
Cypress R., Manitoba, Canada	2000	Sandy loam, 5% OM	49°33′N-99°05′W 374 m asl	Conv. Tillage
Roseworthy, South Australia, Australia	2004	Heavy clay, 3·4% OM	34°32′S-°38°44′E 65 m asl	No-tillage
Roseworthy, South Australia, Australia	2005	Clay loam, 3·1% OM	34°32′S-°38°44′E 65 m asl	No-tillage

minimum air temperature *v*. maximum air temperature ($R^2 = 0{\cdot}60$) and vice versa ($R^2 = 0{\cdot}52$). These regressions were substituted for missing values in the original data set. For missing rainfall events, records from nearby stations were examined; all showed nil values for days with missing data at Roseworthy, except 2 July 2005. An average value of 5 mm rain was inserted into the original data set for that date.

Estimation of soil temperature and soil water potential

The Soil Temperature and Moisture Model (STM^2) was used to estimate soil microclimate conditions (Spokas & Forcella 2009). In the present paper, 20 mm was considered to be a representative seed burial depth for all experimental data sets. The STM^2 uses daily rainfall and minimum and maximum air temperatures, along with soil texture, organic matter and various other simple site descriptors, as input variables to simulate hourly or daily values of soil temperature and soil water potential. The model simulates conditions in ploughed (conventional tillage) and direct-drilled (no-tillage) soils through adjustments in '% shade' of the soil surface (i.e. residue cover) and bulk density. It has been used successfully in both temperate and Mediterranean-type environments around the world (Schutte *et al.* 2008; Masin *et al.* 2012).

Input variables for emergence models

Many definitions of thermal time and hydro-time exist. Moreover, since all the definitions depend on several cardinal parameters (Gardarin *et al.* 2010) many more indices can be defined by making different choices among such parameters. The following indices were used as input variables for model development:

Thermal time (I)

Thermal time (θ_T^I) accumulation was calculated according to Hammer *et al.* (1993):

$$\theta_T^I = \sum_{i=1,n} (T - T_b), \quad \text{if } T_b < T < T_o \tag{1a}$$

$$\theta_T^I = \sum_{i=1,n} (T - T_b)\left(1 - \frac{T_o - T_b}{T_m - T_b}\right), \quad \text{if } T_o < T < T_m \tag{1b}$$

$$\theta_T^I = 0, \quad \text{otherwise} \tag{1c}$$

Equations (1*a*) and (1*b*) are defined for the sub-optimal and supra-optimal thermal ranges, respectively, where the term '$i = 1, n$' represents calendar days, T is the estimated mean daily soil temperature, T_b, T_o and T_m are the base, optimal and maximum temperatures for wild oat seedling emergence, respectively. The following cardinal temperatures values were used: $T_b = 0$ °C (Bullied *et al.* 2003), $T_o = 15$ °C and $T_m = 35$ °C (Sharma *et al.* 1976).

Hydro-time (I)

Hydro-time (θ_H^I) was calculated by Bradford (1990) as

$$\theta_H^I = \sum_{i=1,n} (\Psi - \Psi_b), \quad \text{if } \Psi > \Psi_b \tag{2a}$$

$$\theta_H^I = 0, \quad \text{otherwise} \tag{2b}$$

where Ψ is the estimated mean daily soil water potential and Ψ_b is the base water potential for

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

emergence. A value of $\Psi_b = -0{\cdot}8$ MPa was selected for model development based on numerous iterations of the aforementioned data sets to preliminary hydrothermal-time indices. Values used in the literature for Ψ_b of wild oat germination range from $-0{\cdot}6$ to $-1{\cdot}2$ MPa (Martinson *et al.* 2007).

The following alternative definitions of thermal time (θ_T^{II}) and hydro-time (θ_H^{II}) were adopted (Leguizamón *et al.* 2005) for the calculation of the hydrothermal time index.

Thermal time (II)

Thermal time (θ_T^{II}) accumulation was calculated as

$$\theta_T^{II} = \sum_{i=1,n} (T - T_b), \quad \text{if } T_b < T \tag{3a}$$

$$\theta_T^{II} = 0, \quad \text{otherwise} \tag{3b}$$

Hydro-time (II)

Hydro-time (θ_H^{II}) depicted as a binary variable:

$$\theta_H^{II} = 1, \quad \text{when } \Psi > \Psi_b \tag{4a}$$

$$\theta_H^{II} = 0, \quad \text{when } \Psi < \Psi_b \tag{4b}$$

Hydrothermal time

Hydrothermal time calculation was performed as described by Bradford (2002):

$$\theta_{HT} = \theta_T^{II}\, \theta_H^{II} \tag{5}$$

where θ_T^{II} and θ_H^{II} are defined according to Eqns (3) and (4), respectively.

NLR model

The Weibull model was used as a representative of the NLR approach. The selection of the following function was based on its wide implementation in weed emergence prediction models (Forcella 1998; Schutte *et al.* 2008; Bullied *et al.* 2012):

$$\text{AcEm} = 1 - \exp\left(-\ln(2)\left(\frac{x}{\alpha}\right)^{\beta}\right) \tag{6}$$

In Eqn (6), AcEm is the accumulated emergence (in proportion), x is the applied hydrothermal time index (θ_{HT}) and α and β are the model parameters. An NLR fitting routine was applied for parameter estimation using the Levenberg–Marquardt algorithm of GraphPad Prism Software (GraphPad version 4.0, San Diego, California, USA).

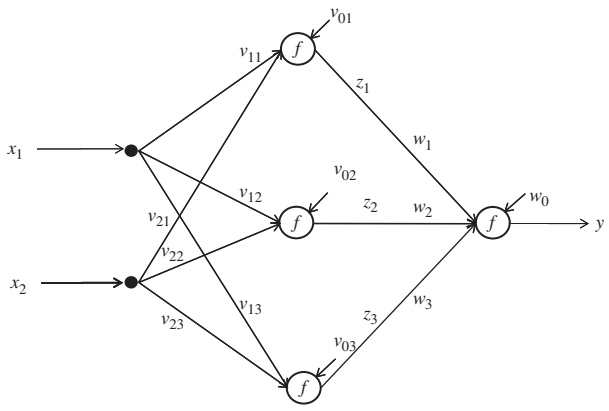


Fig. 1. ANN architecture with three layers, two inputs, one output and three neurons in the hidden layer.

ANNs modelling

Figure 1 shows a three-layer feed-forward ANN. The network has two inputs (x_1, x_2), one output (y) and three neurons in the hidden layer.

Each of the input nodes receives one input (x_1, x_2) and broadcasts it to each one of the hidden neurons. Each hidden neuron computes its activation function and broadcasts its result (z_1, z_2, z_3) to the single output neuron that finally produces the response of the network (y). The output signal of each hidden neuron (z_j) is calculated as

$$z_j = f\left(\sum_{i=1,2} v_{ij} X_i + v_{0j}\right), \quad j = 1, \ldots, 3 \tag{7}$$

while the output of the network is given by

$$y = f\left(\sum_{j=1,3} w_j z_j + w_0\right) \tag{8}$$

In Eqns (7) and (8), the term f(xxxxx) is the activation function of the network, v_{ij} are the weights of the connections between the input and hidden neurons and v_{0j} is the bias on hidden neuron j. Similarly, w_j represents the weights of the connections between the hidden and output neurons and w_0 is the bias of the output neuron.

A bivariate ANN with three neurons in the hidden layer was used as a representative of the neural network approach:

$$\text{AcEm} = \text{ANN}_3\left(\theta_T^I, \theta_H^I\right) \tag{9}$$

In the present work, a feed-forward neural network structure with three layers was adopted (Fig. 1). Several ANNs with different numbers of neurons in the hidden layer were investigated, following the programming

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

guidelines adopted in Chantre *et al.* (2012). Specifically, hyperbolic tangent sigmoid transfer functions were used to model the neurons and the Bayesian Regularization algorithm was selected for training purposes. Besides weights and biases, the effective number of parameters (η), a measure of how many of the parameters of the network are used effectively in reducing the error function, was also calculated.

Importantly, many alternative ANN models could have been generated. In particular, the use of ANNs with θ_{HT} as the only explanatory variable was also explored. Since the corresponding results were comparable to those of the Weibull model (Eqn 6), this version is not included in the present study. Moreover, many ANN models can be generated by increasing the number of neurons in the hidden layer. The optimal number of neurons in the hidden layer in the developed ANN model (Eqn 9) was obtained by trials with increasing numbers of nodes, evaluating typical model performance indicators (test set root mean square error (RMSE) and Akaike's information criteria (AIC)) and the graphical behaviour of the predictions for the test sets (Chantre *et al.* 2012). The Neural Network Toolbox of Matlab (Beale *et al.* 2011) was used for programming the ANNs.

Goodness-of-fit and model performance

In all cases, goodness-of-fit measures were based on the RMSE of the training set. The predictive capability of the developed models was based on the RMSE of the test set:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{10}$$

where y and $\hat{y}$ are the observed and predicted data, respectively. N represents the number of observations.

Comparison between models was performed on AIC (Burnham & Anderson 2002).

Training and test subsets

The available data (Table 1) were divided into training (0·67 of the data set) and testing (0·33 of the data set) subsets. Training and testing data were classified according to Table 2 to represent each location and soil tillage condition adequately.

Table 2. *Training and test sets. NT = no-tillage, CT = conventional tillage*

Training set (0·67 of data set)	Testing set (0·33 of data set)
1. Minnesota 2002 (NT)	9. North Dakota 2003 (NT)
2. Minnesota 2003 (NT)	10. Montana 2002 (CT)
3. North Dakota 2002 (NT)	11. South Australia 2005 (NT)
4. Montana 2000 (CT)	12. Manitoba 2000 (CT)
5. Montana 2001a (CT)	
6. Montana 2001b (CT)	
7. South Australia 2004 (NT)	
8. Manitoba 2000 (NT)	

RESULTS

Model parameters for the univariate NLR and bivariate ANN approaches are shown in Table 3.

The models examined in the present paper provided an adequate fit for the training and test data sets (Table 4). According to the AIC-based model selection technique, the ANN model presented a better performance index than the Weibull model (Table 4). The observed data and the predictions are graphically presented in Fig. 2. Calendar time (Julian days) was used in the abscissas instead of soil microclimatic indices to simultaneously present the predictions of both, univariate and bivariate models.

As observed in Fig. 2, the Weibull model (dotted line) failed to adequately predict wild oat emergence patterns in Montana (RMSE = 0·24), South Australia (RMSE = 0·33) and Manitoba (RMSE = 0·27), but did produce acceptable results for North Dakota (RMSE = 0·13). The univariate NLR model clearly overestimated emergence in South Australia and Manitoba, while a moderate overestimation was observed at the onset of the emergence period in Montana followed by an appreciable underestimation during the rest of the period. A slight underestimation was also registered in North Dakota (Fig. 2). Conversely, the ANN model provided an overall very good fit (Fig. 2, solid line) with an excellent representation in North Dakota (RMSE = 0·03), South Australia (RMSE = 0·06) and Manitoba (RMSE = 0·05) with only a slight overestimation in Montana (RMSE = 0·10). From these results the Weibull model, representative of the NLR approach, was unable to predict accurately both the timing and magnitude of wild oat emergence flushes in the majority of the evaluated locations.

DISCUSSION

The bivariate ANN approach based on thermal time and hydro-time as explanatory variables significantly

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

Table 3. *Parameters estimated for univariate and bivariate models. v_{ij} = connection weights between input-hidden layer neurons, w_j = connection weights between hidden and output layer neurons, v_{0j} = bias on hidden neuron j, w_0 = output neuron bias*

Model	Parameters
Weibull (θ_{HT})	α: 286·3 ± 10·04 β: 2·7 ± 0·34
ANN_3 (θ_T^l, θ_H^l)	v_{11}: −2·67, v_{12}: −17·01, v_{13}: 4·46 v_{21}: 8·58, v_{22}: 32·07, v_{23}: −1·00 w_{11}: 2·27, w_{12}: 0·74, w_{13}: 3·13 v_{01}: 6·60, v_{02}: 5·37, v_{03}: 2·18 w_0: −2·68

Table 4. *Values of the AIC and RMSE for the univariate NLR and bivariate ANN_3 models. m = total number of model parameters, η = number of effective parameters*

Model	m	η	AIC	RMSE train	RMSE test
Weibull (θ_{HT})	2	–	−1·438	0·19	0·24
ANN_3 (θ_T^l, θ_H^l)	13	12	−2·447	0·05	0·07

improved wild oat emergence predictions. These results qualitatively coincide with those presented by Chantre *et al.* (2012), who concluded that univariate NLR and ANN models based on hydrothermal time indices were not able to predict wild oat emergence patterns typical of a semiarid environment as well as bivariate ANN models. The reason for the superiority of the bivariate models might stem from the fact that the emergence process could be described in terms of a germination phase and a post-germination stage, assuming seed dormancy to be negligible. As suggested by other authors (Bradford 2002; Brisson *et al.* 2008), germination depends on both thermal accumulation and moisture conditions, while post-germination elongation of the coleoptile (shoot growth) is mostly driven by thermal time accumulation. The use of hydro-thermal indices as single explanatory variables in univariate models combines both effects in order to minimize the prediction error. Conversely, the adoption of the two independent input variables (i.e. thermal time and hydro-time) weighted differently within a neural network, which allowed a more intimate discrimination of the germination and post-germination processes. These results should alert weed modellers to extrapolate the usage of θ_{HT} from germination to emergence models only with great caution.

Notably, while a bivariate ANN model with 12 effective parameters provided an excellent representation for the regular field emergence data characteristic of temperate environments without considerable restrictive soil moisture limitations, as in the present case, a more complex network with 20 effective parameters was required to represent the irregular and timely distributed emergence observed in a semiarid region (Chantre *et al.* 2012).

An *ad hoc* attempt to use the model developed in the present work to predict wild oat emergence with the weather data used in Chantre *et al.* (2012) showed that field emergence clearly was over-estimated, although it was possible to adequately estimate the onset of the emergence time-window (results not shown). As stated by Grundy *et al.* (2003), predicting the timing of a flush of seedlings is easier than estimating the absolute magnitude of emergence due to fact that the latter is influenced by the seed dormancy status and seed longevity, among other factors.

The differences between wild oat emergence patterns in semiarid conditions (Chantre *et al.* 2012) and those studied in the present work might be attributed mainly to the effect of contrasting seed bank dormancy behaviour due to ecological adaptations to different regional climatic conditions. Seed dormancy is a crucial mechanism to avoid germination under unfavourable environmental conditions for seedling survival, extending seed longevity in the soil (Baskin & Baskin 1998). In highly unpredictable climatic environments, soil temperature and water availability are key environmental factors regulating the dormancy status of weed seed banks (Bouwmeester 1990). Indeed, for wild oat specifically, the interplay of temperature and moisture on dormancy relief is complex. Dormancy relief is fastest and greatest at high temperatures and low seed moisture (e.g. 40 °C, 10%). However, at low temperatures (e.g. 20 °C), high seed moisture (18%) allows maximum dormancy relief (Foley 1994).

Evidently, wild oat has developed specific ecological adaptations to different regional climatic conditions with respect to dormancy, perhaps *via* selection of genes at three loci as described by Fennimore *et al.* (1999). In temperate or Mediterranean environments with either non-severe soil moisture availability limitations or seasonal limitations that are predictable during the emergence period, regular and concentrated cumulative emergence curves are observed

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

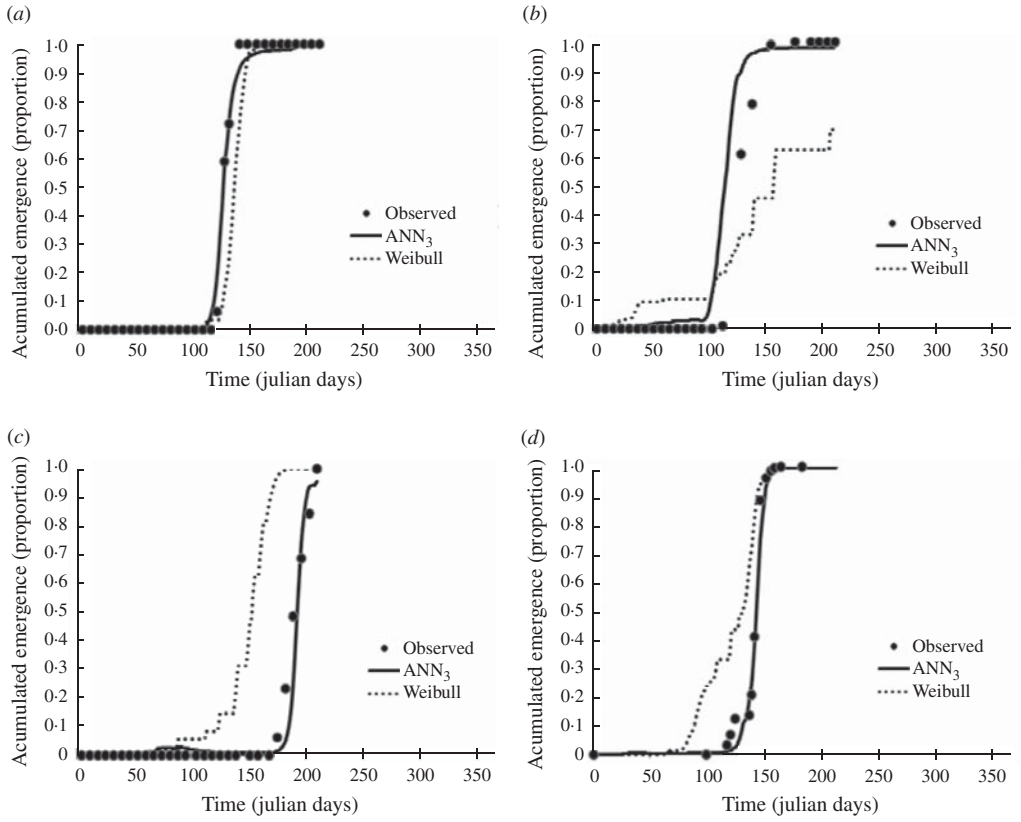


Fig. 2. Observed *v.* predicted wild oat cumulative emergence curves for the test set using Weibull (θ_{HT}) (dotted line) and ANN_3 (θ_T^I, θ_H^I) (solid line) models. (*a*) North Dakota 2003 (NT), (*b*) Montana 2002 (CT), (*c*) South Australia 2005 (NT) and (*d*) Manitoba 2000 (CT). NT = no-tillage, CT = conventional tillage.

(Chauhan *et al.* 2006; Page *et al.* 2006; Martinson *et al.* 2007; the present paper). Conversely, in semiarid regions typical emergence patterns are very irregular with multiple cohorts distributed within a considerably large time-span (Chantre *et al.* 2012). For the latter case, such behaviour could be attributed to a highly unpredictable annual precipitation regime and also to the effect of a heterogeneous maternal environment during seed maturation (López & Vigna 1991). Maternal effects related to climatic differences during seed development are expected to influence greatly the dormancy behaviour of a given species population or ecotype (Baskin & Baskin 1998), thus contributing to potentially unpredictable variability of the observed emergence data (Grundy *et al.* 2003). As suggested by Peters (1982), the effect of a variable environment during seed maturation in the panicle is expected to affect long-term dynamics of wild oat populations.

These ecological adaptations prevent, at the moment, the development of 'universal' weed emergence predictive models. A deeper understanding of the factors that regulate the sub-processes that take place before and during emergence (dormancy release, germination and initial shoot growth) should contribute to the development of such universal predictors.

Overall, it might be concluded that bivariate ANN models constitute a relatively simple modelling framework and provide adequate representation of wild oat field emergence when developed for specific environmental conditions. Furthermore, from a practical agronomic perspective, these features encourage the use of ANNs as emergence predictors within weed management decision support systems, but perhaps more importantly, ANN modelling contributed to the understanding of the differences in emergence behaviour among populations of weed species (wild oat, in this case). In addition, it helped to rationalize why developing universal emergence models for wild oat may not be possible. The latter result is important in a practical sense and points out the importance of understanding weed species populations on a regional basis and developing or modifying management regimes accordingly.

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

This research was partially supported by grants from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET PIP No 11220100100222) and Universidad Nacional del Sur (PGI 24/A157) to G.R.C., A.M.B. and M.R.S. and by grants from FEDER funds and the Spanish Ministry of Economy and Competitiveness (AGL2009-7883) to J.L.G.-A.

REFERENCES

Alvarez, R. (2009). Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *European Journal of Agronomy* **30**, 70–77.

Baskin, C. C. & Baskin, J. M. (1998). *Seeds: Ecology, Biogeography and Evolution of Dormancy and Germination*. San Diego, CA: Academic Press.

Batlla, D. & Benech-Arnold, R. L. (2007). Predicting changes in dormancy level in weed seed soil bank: implications for weed management. *Crop Protection* **26**, 189–197.

Beale, M. H., Hagan, M. T. & Demuth, H. B. (2011). *Neural Network Toolbox, User's Guide, MATLAB*. Natick, MA: The MathWorks Inc.

Bouwmeester, H. J. (1990). *The effect of environmental conditions on the seasonal dormancy pattern and germination of weed seeds*. Ph.D. Thesis, Wageningen Agricultural University, The Netherlands.

Bradford, K. J. (1990). A water relations analysis of seed germination rates. *Plant Physiology* **94**, 840–849.

Bradford, K. J. (2002). Applications of hydrothermal time to quantifying and modelling seed germination and dormancy. *Weed Science* **50**, 248–260.

Brisson, N., Launay, M., Mary, B. & Beaudoin, N. (2008). *Conceptual Basis, Formalizations and Parameterization of the STICS Crop Model*. Paris, France: Quae Editions.

Bullied, W. J., Marginet, A. M. & Van Acker, R. C. (2003). Conventional- and conservation-tillage systems influence emergence periodicity of annual weed species in canola. *Weed Science* **51**, 886–897.

Bullied, W. J., Van Acker, R. C. & Bullock, P. R. (2012). Hydrothermal modelling of seedling emergence timing across topography and soil depth. *Agronomy Journal* **104**, 423–436.

Burnham, K. P. & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York: Springer-Verlag.

Cao, R., Francisco-Fernández, M., Anand, A., Bastida, F. & González-Andújar, J. L. (2011). Computing statistical indices for hydrothermal times using weed emergence data. *Journal of Agricultural Science, Cambridge* **149**, 701–712.

Chantre, G. R., Blanco, A. M., Lodovichi, M. V., Bandoni, J. A., Sabbatini, M. R., López, R. L., Vigna, M. R. & Gigón, R. (2012). Modelling *Avena fatua* seedling emergence dynamics: an artificial neural network approach. *Computers and Electronics in Agriculture* **88**, 95–102.

Chauhan, B. S., Gill, G. & Preston, C. (2006). Seedling recruitment pattern and depth of recruitment of 10 weed species in minimum tillage and no-till seeding systems. *Weed Science* **54**, 658–668.

Colbach, N. & Mézière, D. (2013). Using a sensitivity analysis of a weed dynamics model to develop sustainable cropping systems. I. Annual interactions between crop management techniques and biophysical field state variables. *Journal of Agricultural Science, Cambridge* **151**, 229–246.

Colbach, N., Chauvel, B., Darmency, H. & Tricault, Y. (2011). Sensitivity of weed emergence and dynamics to life-traits of annual spring-emerging weeds in contrasting cropping systems, using weed beet (*Beta vulgaris* ssp. vulgaris) as an example. *Journal of Agricultural Science, Cambridge* **149**, 679–700.

Cousens, R. D., Weaver, S. E., Martin, T. D., Blair, A. M. & Wilson, J. (1991). Dynamics of competition between wild oats (*Avena fatua* L.) and winter cereals. *Weed Research* **39**, 203–210.

Foley, M. E. (1994). Temperature and water status of seed affect after ripening in wild oat (*Avena fatua*). *Weed Science* **42**, 200–204.

Fennimore, S. A., Nyquist, W. E., Shaner, G. E., Doerge, R. W. & Foley, M. E. (1999). A genetic model and molecular markers for wild oat (*Avena fatua* L.) seed dormancy. *Theoretical and Applied Genetics* **99**, 711–718.

Forcella, F. (1998). Real-time assessment of seed dormancy and seedling growth for weed management. *Seed Science Research* **8**, 201–209.

Forcella, F., Benech-Arnold, R. L., Sánchez, R. & Ghersa, C. M. (2000). Modelling seedling emergence. *Field Crops Research* **67**, 123–139.

Gardarin, A., Durr, C. & Colbach, N. (2012). Modelling the dynamics and emergence of a multispecies weed seed bank with species traits. *Ecological Modelling* **240**, 123–138.

Gardarin, A., Guillemin, J. P., Munier-Jolain, N. M. & Colbach, N. (2010). Estimation of key parameters for weed population dynamics models: Base temperature and base water potential for germination. *European Journal of Agronomy* **32**, 162–168.

González-Andújar, J. L., Forcella, F., Kegode, G., Gallagher, R. & Van Acker, R. (2001). Modelizacion de la emergencia de plantulas de Avena loca (*Avena fatua* L.) usando tiempo hidrotermal. In *Congreso 2001 de la Sociedad Española de Malherbologia* (Eds J. A. Boto, J. L. Villarias & J. C. Fernandez), pp. 243–246. Leon, Spain: Universidad de León.

Grundy, A. C. (2003). Predicting weed emergence: a review of approaches and future challenges. *Weed Research* **43**, 1–11.

Grundy, A. C., Peters, N. C. B., Rasmussen, I. A., Hartmann, K. A., Sattin, M., Andersson, L., Mead, A., Murdoch, A. J. & Forcella, F. (2003) Emergence of *Chenopodium album* and *Stellaria media* of different origins under different climatic conditions. *Weed Research* **43**, 163–176.

Haj Seyed Hadi, M. R. & González-Andújar, J. L. (2009). Comparison of fitting weed seedling emergence models with nonlinear regression and genetic algorithm. *Computers and Electronics in Agriculture* **65**, 19–25.

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press

Hammer, G. L., Carberry, P. S. & Muchow, R. C. (1993). Modelling genotypic and environmental control of leaf area dynamics in grain sorghum. I. Whole plant level. *Field Crops Research* **33**, 293–310.

Izquierdo, J., González-Andújar, J. L., Bastida, F., Lezaún, J. A. & Sánchez Del Arco, M. J. (2009). A thermal time model to predict corn poppy (*Papaver rhoeas*) emergence in cereal fields. *Weed Science* **57**, 660–664.

Leguizamón, E. S., Fernandez-Quintanilla, C., Barroso, J. & Gonzalez-Andujar, J. L. (2005). Using thermal and hydrothermal time to model seedling emergence of *Avena sterilis* ssp. *ludoviciana* in Spain. *Weed Research* **45**, 149–156.

Lek, S. & Guégan, J. F. (1999). Artificial neural networks as a tool in ecological modelling: an introduction. *Ecological Modelling* **120**, 65–73.

López, R. L. & Vigna, M. R. (1991). Patrón de emergencia de *Avena fatua* L. en el sudoeste de Buenos Aires. In *Actas X Reunión Nacional CAPERAS* (Ed. Universidad Nacional del Sur), pp. 135–136. Bahía Blanca, Argentina: Universidad Nacional del Sur.

Martinson, K., Durgan, D., Forcella, F., Wiersma, J., Spokas, K. & Archer, D. (2007). An emergence model for wild oat (*Avena fatua*). *Weed Science* **55**, 584–591.

Masin, R., Loddo, D., Benvenuti, S., Otto, S. & Zanin, G. (2012). Modelling weed emergence in Italian maize fields. *Weed Science* **60**, 254–259.

Mickelson, J. A. & Grey, W. E. (2006). Effect of soil water content on wild oat (*Avena fatua*) seed mortality and seedling emergence. *Weed Science* **54**, 255–262.

Moschini, R. C., López, R. L., Vigna, M. R. & Damiano, F. (2009). Modelos basados en tiempo térmico e hidrotérmico para predecir la emergencia de Avena fatua en lotes con y sin labranza estival, en Argentina. In *Herbología e Biodiversidade numa Agricultura Sustentável* (Eds E. de Sousa, I. M. Calha, I. Moreira, L. Rodrigues, J. Portugal & T. Vasconcelos), pp. 239–242. Lisboa, Portugal: ISA Press.

Onofri, A., Gresta, F. & Tei, F. (2010). A new method for the analysis of germination and emergence data of weed species. *Weed Research* **50**, 187–198.

Onofri, A., Mesgaran, M. B., Tei, F. & Cousens, R. D. (2011). The cure model: an improvement way to describe seed germination. *Weed Research* **51**, 516–524.

Page, E. R., Gallagher, R. S., Kemanian, A. R., Zhang, H. & Fuerst, E. P. (2006). Modelling site-specific wild oat (*Avena fatua*) emergence across a variable landscape. *Weed Science* **54**, 838–846.

Peters, N. C. B. (1982). The dormancy of wild oat seed (*Avena fatua* L.) from plants grown under various temperature and soil moisture conditions. *Weed Research* **22**, 205–212.

Schutte, B. J., Regnier, E. E., Harrison, S. K., Schmoll, J. T., Spokas, K. & Forcella, F. (2008). A hydrothermal seedling emergence model for giant ragweed (*Ambrosia trifida*). *Weed Science* **56**, 555–560.

Sharma, M. P., Mcbeath, D. K. & Vanden Born, W. H. (1976). Studies on the biology of wild oats. I. Dormancy, germination and emergence. *Canadian Journal of Plant Science* **56**, 611–618.

Spokas, K. & Forcella, F. (2009). Software tools for weed seed germination modelling. *Weed Science* **57**, 216–227.

https://doi.org/10.1017/S0021859612001098 Published online by Cambridge University Press