Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

A flexible and practical approach for real-time weed emergence prediction based on Artificial Neural Networks

Guillermo R. Chantre [a,b,*], *Mario R. Vigna* [c], *Juan P. Renzi* [a,d], *Aníbal M. Blanco* [e]

[a] *Departamento de Agronomía, Universidad Nacional Del Sur, San Andrés 800, 8000 Bahía Blanca, Argentina*
[b] *Centro de Recursos Naturales Renovables de La Zona Semiárida (CERZOS), Universidad Nacional del Sur, CONICET, 8000 Bahía Blanca, Argentina*
[c] *EEA INTA Bordenave, Bordenave, Buenos Aires 8187, Argentina*
[d] *Instituto Nacional de Tecnología Agropecuaria, 8142 Hilario Ascasubi, Argentina*
[e] *Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur, CONICET, 8000 Bahía Blanca, Argentina*

ARTICLE INFO

Article history:
Received 13 July 2017
Received in revised form
26 March 2018
Accepted 29 March 2018

Keywords:
Integrated Weed Management Support Systems
Semiarid region
Soft computing
Artificial intelligence

Most popular emergence prediction models require species-specific population-based parameters to modulate thermal/hydrothermal accumulation. Such parameters are frequently unknown and difficult to estimate. Moreover, such models also rely on hardly available and difficult to estimate soil site-specific microclimate conditions, which in turn depend on soil heterogeneity at a field spatial level. On the other hand, modern agriculture benefits from easily available real-time information, in particular on-line meteorological data generated by forecasts and automatic local weather stations. In this context, Artificial Neural Networks (ANN) provide a flexible option for the development of prediction models, especially to study species which show a highly distributed emergence pattern along the year. In this work, an ANN approach based on easily obtainable meteorological data (daily minimum and maximum temperatures; daily precipitation) is proposed for weed emergence prediction. Relative Daily Emergence (RDE), expressed as a proportion of the total emergence, was the adopted output variable. Field emergence data recorded on a weekly basis were used to generate RDE patterns through linear interpolation. Results for three study cases from the Semiarid Pampean Region of Argentina (*Lolium multiflorum*, *Avena fatua* and *Vicia villosa*), which show irregular and time-distributed field emergence patterns, are reported. In all cases, ANN model selection was based on the Root Mean Square Error of the test set which showed better consistency than other typical Information Theory performance metrics. The combination of large ANN with a Bayesian Regularization Algorithm generated satisfactory estimations based on the RMSE values for independent Cumulative Emergence data.

* *Corresponding author*. Departamento de Agronomía, Universidad Nacional del Sur, Av. Colón 80, Bahía Blanca 8000, Argentina. Fax: +54 291 4595127.
E-mail address: gchantre@criba.edu.ar (G.R. Chantre).

https://doi.org/10.1016/j.biosystemseng.2018.03.014

1. Introduction

Sigmoidal Regression Models (SRM) based on thermal or hydrothermal-time, adequately represent weed emergence patterns under certain ecological environments, such as in temperate regions with non-severe soil moisture restrictions (Gonzalez-Andujar, Chantre, Morvillo, Blanco, & Forcella, 2016). Conversely, weed emergence prediction under arid or semiarid environments with highly variable soil thermal and moisture conditions represents a challenging task due to the occurrence of irregular, time-distributed field patterns. Despite this fact, some SRM models have been successfully developed and validated (Royo-Esnal, García, Torra, Forcella, & Recasens, 2015).

In general, weed emergence modelling approaches require the implementation of thermal or hydrothermal indexes. Such indexes are based on soil microclimatic variables (soil temperature and water potential) which are in turn dependent on many site-specific variables (e.g. soil texture, surface cover, seed burial depth). Indexes construction depends on: (i) the assumption that emergence rates are proportional to the amount by which soil temperature and soil water potential exceed a given threshold value (Bradford, 2002); (ii) the estimation of site-specific soil microclimate variables using specific software (e.g. Soil Temperature and Moisture Model) (Spokas & Forcella, 2009). In addition, species-specific thresholds or cardinal parameters (i.e. base temperature and base water potential for seed germination/emergence) are generally obtained under laboratory-controlled conditions based on certain statistical principles (Bradford, 2002): (i) independence between threshold parameters and soil microclimatic variables (soil temperature and water potential); and (ii) a normal distribution of thermal/hydrothermal-time among individuals of a given seed population. Alternatively, these cardinal parameters can also be obtained through field trials and modelling (Royo-Esnal, Torra, Conesa, Forcella, & Recasens, 2010; Royo-Esnal et al., 2015).

In the Semiarid Pampean Region of Argentina (SPRA), *Avena fatua* L. and *Lolium multiflorum* Lam. are among the most conspicuous and invasive weed species in winter cereal crops. Both species present time-distributed multiple emergence cohorts *per* year, thus hindering the implementation of traditional SRM. In the southern area of the region under study, *Vicia villosa* Roth. is considered a problematic volunteer weed in winter cereals crop rotations, showing a “double-safety” seed dormancy mechanism (physical + physiological dormancy) (Renzi, Chantre, & Cantamutto, 2014) which regulates field emergence flushes. Recently, Renzi, Chantre, and Cantamutto (2018) made a considerable effort to develop and validate a field emergence model for *V. villosa* by integrating: (i) physical dormancy release dynamics, (ii) physiological dormancy release and germination thermal requirements, (iii) hydro-time requirements for germination, and (iv) pre-emergence growth.

In this regard, mechanistic or population-based models should be considered a valuable tool for the description of basic ecophysiological processes underlying seedling emergence (Boddy, Bradford, & Fischer, 2012; Chantre, Batlla, Sabbatini, & Orioli, 2009; Colbach & Mézière, 2013). Although such models are desirable from an explanatory biological process-based framework, their development and validation are usually very time consuming. As suggested by Grundy (2003), such models lack of the simplicity and flexibility that would be required for practical Integrated Weed Management Support Systems (IWMSS), which weather-based models do offer.

Based on previously exposed statements, it is clear that current weed emergence modelling is in some extent driven by a *precise-and-deterministic* view of a highly uncertain and complex ecological problem resulting from the interaction of weed biology and soil variables. In addition, bioecological limitations arise as seed dormancy processes and germination requirements of various weed species remain to be elucidated (Batlla & Benech-Arnold, 2010) in order to bridge the ecological knowledge gap for accurate weed emergence prediction.

Recently, Gonzalez-Andujar et al. (2016) highlighted the importance of a new generation of modelling approaches based on Soft Computing Techniques (SCT) to tackle the limitations of conventional nonlinear regression models. It was concluded in that work that SCT provide higher flexibility as well as better predictive accuracy in many cases.

SCT (unlike conventional or hard computing) are capable of dealing with complex biological systems because do not require strict mathematical definitions, thus providing a better rapport with reality (Das, Kumar, Das, & Burnwal, 2013). Among them, for example, Artificial Neural Networks (ANN) as modelling framework and Genetic Algorithms (GA) as optimization engines, have been applied in most fields of science and technology (Gen & Cheng, 2000; Paliwal & Kumar, 2009), and also in many fields of agriculture (Matsumura, Gaitan, Sugimoto, Cannon, & Hsieh, 2015; Pi et al., 2015). For example, Pi et al. (2015) proposed an ANN-GA combined model for the prediction of optimal sowing time and cultivars selection across different regions of China working with the grass *Poa pratensis* L.

ANN and GA have also been applied in weed research (Burgos-Artizzu, Ribeiro, Tellaeche, Pajares, & Fernández-Quintanilla, 2010; Burks, Shearer, Heath, & Donohue, 2005; Dyrmann, Karstoft, & Midtiby, 2016). However, soft computing modelling for weed emergence prediction remains largely unexplored (Gonzalez-Andujar et al., 2016). Haj Seyed Hadi and Gonzalez-Andujar (2009) compared GA with nonlinear regression for fitting emergence data of six weed species of Spanish cereal crops showing that the former often provided a better fit due to a higher capability to deal with ill-defined optimisation issues. Blanco et al. (2014) proposed a GA approach for *A. fatua* emergence prediction based on the disaggregation of dormancy release and germination/pre-emergence growth processes. Previously, Chantre et al. (2012) developed an ANN model using both thermal-time and hydro-time variables as independent explanatory variables for the biosystem under study in the SPRA. The same approach proved adequate for *A. fatua* emergence prediction in different temperate regions of the United States, Canada and south Australia (Chantre et al., 2012, 2014).

In previously mentioned studies, soft computing based models have proven to yield better prediction capacity than classical univariate hydrothermal-time based nonlinear regression models. However, some caveats were detected by