

APLICACIÓN DE INTELIGENCIA COMPUTACIONAL Y COMPUTACIÓN DE ALTO DESEMPEÑO EN EL DESARROLLO DE UN MODELO DE PREDICCIÓN DE LAS CONDICIONES PREDISPONENTES AL QUEMADO DEL ARROZ (PYRICULARIA ORYZAE)

Miriam Asselborn⁽⁴⁾, Julian Escalante⁽¹⁾, Mariela Lopresti⁽²⁾, Natalia Miranda⁽²⁾, Esteban Schab^(1,3), Karina Cedaro⁽¹⁾, Pablo Fontanini⁽⁶⁾, Malvina Martinez⁽⁵⁾, Carlos Casanova^(1,3), Virginia Pedraza⁽⁴⁾, Fabiana Piccoli^(1,2)

⁽¹⁾ Universidad Autónoma de Entre Ríos, Facultad de Ciencia y Tecnología, Concepción del Uruguay.

⁽²⁾ LIDIC- Univ. Nacional de San Luis, San Luis.

⁽³⁾ Univ. Tecnológica Nacional, Facultad Regional Concepción del Uruguay, Concepción del Uruguay.

⁽⁴⁾ INTA, Estación Experimental Agropecuaria Concepción del Uruguay, Concepción del Uruguay.

⁽⁵⁾ Instituto de Clima y Agua. CIRN. INTA Castelar, Buenos Aires.

⁽⁶⁾ Bolsa de Cereales de Entre Ríos.

Argentina

{omlopres,ncmiran,mpiccoli}@unsl.edu.ar

{schabe,casanovac}@frcu.utn.edu.ar

RESUMEN

Un sistema complejo o crítico con toma de decisiones se caracteriza por la imposibilidad de reproducir para su estudio un escenario sin consecuencias reales, o cuando su resolución implica gran cantidad de recursos para obtener resultados en un tiempo prudencial. La complejidad puede darse por las características del problema o por la cantidad de datos con los que se trabaja. Tomar decisiones en estos contextos debe conjugar dos atributos usualmente contrapuestos: calidad y velocidad. En este trabajo proponemos una línea de investigación enfocada en analíticas, principalmente las prescriptivas, capaces de determinar acciones a ser ejecutadas en el momento (decisiones operativas) o en el futuro (decisiones tácticas: corto y mediano plazo, decisiones estratégicas: largo plazo) para lograr un objetivo deseado. A esta línea se suman investigaciones en Inteligencia Computacional y Computación de Alto Desempeño con el fin de obtener, de forma colaborativa, calidad y velocidad en las decisiones.

Palabras clave: Inteligencia Computacional. Analíticas. Big Data. Computación de Alto Desempeño.

CONTEXTO

Esta propuesta de trabajo se lleva a cabo dentro de los proyectos de investigación: “Inteligencia Computacional aplicada al desarrollo de un modelo Eficiente de predicción de las condiciones predisponentes al Quemado del Arroz (*Pyricularia oryzae*) en Argentina” (PICTO-UADER-UNER-00020, UADER e INTA) y “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” (LIDIC, UNSL).

1. INTRODUCCIÓN

El modelado computacional es una gran herramienta para poder generar nuevos conocimientos en diferentes áreas de la ciencia como la biología, física y agronomía [11, 18,24] entre otras. El uso de herramientas y técnicas computacionales permite llevar a cabo el estudio de un sistema y el análisis de los posibles resultados para tomar buenas decisiones. Entre ellas se encuentran la inteligencia computacional, la simulación y la computación de alto desempeño [4,25].

Desarrollar sistemas complejos para la toma de decisiones basados en técnicas de simulación e inteligencia artificial es una de las prácticas más difundidas a la hora de resolver computacionalmente problemas muy grandes, ya sea por su complejidad o por la cantidad de datos involucrados, generalmente obtenidos de la realidad en forma off-line o en tiempo real. Los sistemas reales complejos, generalmente, no son deterministas ni dicotómicos y requieren, frecuentemente de gran cantidad de datos, imposibles de reconocer, procesar y comprender al mismo tiempo por un humano [9,27,28]. Un modelo computacional de estos sistemas debe maximizar su utilidad, y para ello se debe lograr un balance apropiado entre: complejidad, credibilidad e incertidumbre [19]. Uno de los enfoques para lograrlo es el de la Inteligencia Computacional (IC), también referida como Soft Computing, la cual se inspira en la naturaleza para desarrollar sistemas inteligentes basados en computadoras. La estrategia de la IC es explotar la tolerancia a la incertidumbre, aproximación y falta de información para lograr eficiencia, robustez y soluciones de bajo costo. Ejemplos de ella son las Redes Neuronales Artificiales, los Sistemas Difusos y la Computación Evolutiva [7]. Todos estos métodos pueden ser utilizados juntos y complementariamente para resolver un problema [23]. En un contexto de toma de decisiones, la IC tiende a soportar el proceso aprovechando la imprecisión e incertidumbre del mismo para brindar “soluciones más satisfactorias”. Las decisiones se derivan de analíticas, las cuales generan el conocimiento necesario según los datos disponibles. Existen distintos tipos y niveles de analíticas, el nivel depende del grado de automatización en el proceso. Su diseño no es una tarea sencilla, tampoco su implementación computacional, ya que se deben considerar, por ejemplo, aspectos como la toma de decisiones multicriterio y multipersona, la fusión de información de

diferentes fuentes y naturalezas, el análisis de datos e información, etc.

Por todo lo expuesto, implementar una solución computacional para un sistema complejo mediante un modelo que soporte la toma de decisiones demanda gran cantidad de recursos a fin de obtener resultados en un tiempo adecuado. La inclusión de técnicas y estrategias de computación de alto desempeño es mandatoria, principalmente la Computación Paralela [8,14,17], más teniendo en cuenta la naturaleza de la IC y las características de los recursos informáticos actuales [20,21].

Un sistema real complejo es el relacionado con la producción de arroz. El arroz es el alimento básico más importante a nivel mundial, después del trigo, en Argentina tiene importancia regional, se cultiva bajo riego por inundación, en las provincias de Entre Ríos, Corrientes, Santa Fe, Chaco y Formosa [22]. El Quemado del Arroz (QA), causado por el hongo *Pyricularia oryzae*, es la enfermedad más importante del cultivo a escala mundial [3,10]. En Argentina, es la enfermedad más temida, su aparición es esporádica, pero puede causar pérdidas totales a nivel de lote [16].

En cuanto al manejo del QA, si bien la resistencia genética es la principal herramienta, las variedades más difundidas en Argentina son susceptibles al patógeno. El control químico, mediante aplicaciones de fungicidas registrados para el QA, durante la floración o principio de llenado de granos, ayuda a disminuir los daños. No obstante, el patógeno posee una gran facilidad para desarrollar resistencias a los fungicidas utilizados para su tratamiento [5].

Los síntomas del QA pueden aparecer en cualquier etapa de desarrollo. La intensidad varía con el ambiente y con la susceptibilidad del cultivar. En Argentina, el síntoma más temido es la necrosis del raquis de panoja, afectando directamente al rendimiento. Según las condiciones naturales (suelo, clima, entre otras) y la intensidad de los monitoreos de los lotes, la

necrosis del raquis de panoja puede ser el primer síntoma de la enfermedad detectado por el productor/asesor en el campo y, en general, suele ser demasiado tarde como para una aplicación exitosa de fungicidas.

Las condiciones meteorológicas predisponentes al QA han sido ampliamente estudiadas a nivel mundial. Los factores más importantes en el desarrollo de la enfermedad son la temperatura, la humedad relativa, la nubosidad, y las horas de mojado de hoja [6]. En condiciones naturales de campo, las prácticas de manejo del cultivo, como la elección del cultivar, el riego y la fertilización nitrogenada, pueden alterar la precisión de las predicciones de aparición de la enfermedad basadas en datos meteorológicos. La relación entre factores zonales y locales, y la intensidad de enfermedad, están siendo estudiadas para el desarrollo de modelos de predicción en Entre Ríos [16].

Por todo lo expuesto, y según la bibliografía existente, el desarrollo del QA constituye un sistema candidato a ser estudiado mediante métodos computacionales de distinta índole. Se han reportado al menos 52 modelos predictivos, de los cuales sólo ocho se encuentran en uso actualmente en todo el mundo, y sólo 4 han sufrido posteriores revisiones [6]. No obstante, nuevos modelos de predicción surgen debido a dificultades que los existentes exhiben, sobre todo relativos a la calidad e integridad de los datos, y al alcance geográfico de los mismos. Recientemente, se ha estudiado la efectividad de algunos métodos provenientes del *machine learning* (por ejemplo, generación de reglas M5 y Redes Neuronales Artificiales LSTM) para la predicción de la aparición del QA, con resultados competitivos respecto a modelos clásicos como el de Yoshino y WARN [2,12,15,26]. Sin embargo, existen aún muchas técnicas avanzadas de la inteligencia computacional que no se han probado en este problema y que pueden resultar superadoras respecto de las ya analizadas. Entre ellas se

encuentran los autómatas celulares como modelos, la lógica difusa para la generación de reglas y otros tipos de arquitecturas de redes neuronales artificiales, así como las posibles hibridaciones entre tales técnicas.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Accionar ante la real o posible presencia de QA en los cultivos, puede determinarse que es un sistema complejo y apto para el desarrollo de herramientas que permitirán analizar la interacción de distintos parámetros en un corto tiempo, y generar herramientas para el productor agropecuario que le permita tomar decisiones y actuar en forma predictiva adelantándose a los hechos o ante los primeros indicios de la enfermedad. Asimismo, se evitarían aplicaciones preventivas innecesarias de fungicidas, con beneficios al medioambiente y a la economía de la producción.

Desarrollar un sistema computacional para la detección temprana del QA implica diferentes líneas de investigación, podemos destacar a:

- Modelos de inteligencia computacional.
- Procesamiento de imágenes y realidad aumentada.
- Tecnologías del Lenguaje Humano y Visualización de la información.
- Computación de alto desempeño, a aplicar en cada una de las otras líneas y en la mayor parte del proceso.

Cada una de estas líneas aportarán al desarrollo de una herramienta de software con emisión de alertas para un accionar preventivo, de manera inteligente, ante la presencia de condiciones predisponentes para el desarrollo de QA.

Con toda la investigación que implica y el desarrollo final, no solamente se introduciría valor agregado al sistema productivo arrocero, permitiendo a los productores un accionar predictivo y proactivo, resultando en una reducción de los costos y, consecuentemente,

una mejora sensible en la seguridad de la cosecha, sino también generar conocimiento relacionado con el comportamiento del QA y las formas de tratamiento, como así también respecto a la aplicación del modelo en problemáticas similares, ya sea en el mismo dominio u otro distinto.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se está trabajando en la modelización del sistema mediante Autómatas Celulares y Agentes de Aprendizaje. Para ellos se necesita el estudio previo del comportamiento de los cultivos y del QA, como así también de los parámetros que influyen en su desarrollo. Inicialmente, se está trabajando de manera incremental respecto a los parámetros. Además, se están considerando otras herramientas de IC para incluir. En todos los casos se plantean las soluciones mediante la aplicación de HPC en GPU.

Como uno de los objetivos es lograr soluciones paralelas portables, de costo predecible, capaz de explotar las ventajas de modernos ambientes HPC a través de herramientas y “frameworks de computación” de alto nivel [14], los primeros desarrollos se están haciendo en GPU Nvidia con CUDA [13], y se prevé el uso de otras tecnologías como OpenCL[1].

4. FORMACIÓN DE RECURSOS HUMANOS

Los resultados esperados respecto a la formación de recursos humanos son el desarrollo de 2 tesis de doctorado y de varias tesinas de grado en las universidades intervinientes.

5. BIBLIOGRAFÍA

- [1] M. Breyer, G. Daiss and D. Pflüger. “Performance - Portable Distributed k-Nearest Neighbors Using Locality - Sensitive Hashing and SYCL”. Int. Workshop on OpenCL. ISBN 9781450390330. Association for Computing Machinery. 2021
- [2] Bregaglio, S., Titone, P., Cappellia, G, Tamborini, L., Mongiano, G., Confalonieri, R.. Coupling a generic disease model to the WARM rice simulator to assess leaf and panicle blast impacts in a temperate climate. *Eur J Agron.* 76:107–17. 2016.
- [3] Carmona, M. Tizón del arroz. Herbario. UBA:http://herbariofitopatologia.agro.uba.ar/?page_id=5571. 2021.
- [4] Huerta, E.A., Khan, A., Davis, E., Bushell, C., Gropp, W., Katz, D., Kindratenko, V., Koric, S., Kramer, W., McGinty, B., McHenry, K., Saxton, A.. Convergence of artificial intelligence and high performance computing on NSF-supported cyberinfrastructure. *J Big Data* 7, 88. Springer. 2020.
- [5] Ishii, H., Hollomon, D.. *Fungicide Resistance in Plant Pathogens : Principles and a Guide to Practical Management*. Springer. ISBN 978-4431556411. 2015.
- [6] Katsantonis, D., Kadoglidou, K., Dramalis, C., Puigdollers, P. Rice blast forecasting models and their practical value: a review. *Phytopathologia Mediterranea*, 56, 2. ISSN 0031-9465. Pp 187–216. 2017.
- [7] Keller, J., Liu, D. Fogel, D.. *Fundamentals of computational intelligence: Neural networks, fuzzy systems, and evolutionary computation*. John Wiley & Sons. ISBN 9781119214342. 2016.
- [8] Kirk, D., Hwu, W.. *Programming Massively Parallel Processors: A Hands-on Approach*. Morgan Kaufmann. 3r ed. ISBN 978012811987. 2016
- [9] Klir, G.. "Uncertainty and information: foundations of generalized information theory." *Kybernetes* . ISSN: 0368-492X, 35 (7-8). Pp. 1297-1299. 2006.
- [10] Li, S., Zhang, X., Wang, W., Guo, X., Wu, Z., Du, W., ... & Xia, L.. Expanding the scope of CRISPR/Cpf1-mediated genome editing in rice. *Molecular Plant*, 11(7). Pp 995-998. 2018.
- [11] Liu, J., Zheng, Q., Mu, X., Zuo, Y., Xu, B., Jin, Y., Li, Y.. Automated tumor proportion score analysis for PD-L1 (22C3) expression

- in lung squamous cell carcinoma. *Scientific reports*, 11(1). Pp 1-9. 2021.
- [12] Nettleton, D. F., Katsantonis, D., Kalaitzidis, A., Sarafijanovic-Djukic, N., Puigdollers, P., Confalonieri, R. Predicting rice blast disease: machine learning versus process-based models. *BMC bioinformatics*. vol. 20, no 1, p. 1-16. 2019.
- [13] Nvidia. "CUDA C++ Programming Guide". https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf. 2023.
- [14] Pacheco, P., Malensek, M. *An Introduction to Parallel Programming*. Morgan Kaufmann. 2 Ed. ISBN 0128046058. 2021.
- [15] Pagani, V., Guarneri, T., Busetto, L., Ranghetti, L., Boschetti, M., Movedi, E., Campos-Taberner, M., Garcia-Haro, F., Katsantonis, D., Stavrakoudis, D., Ricciardelli, E., Romano, F., Holecz, F., Collivignarelli, F., Granell, C., Casteleyn, S., Confalonieri, R. A high-resolution, integrated system for rice yield forecasting at district level. *Agricultural Systems*, 168. Pp 181-190. ISSN 0308-521X. 2019.
- [16] Pedraza, M.V.; Liberman, C.A.; Nuñez Bordoy, E.; Sanso, C.; Panozzo, J.; Asselborn, M.N. Control químico del Quemado del Arroz (*Pyricularia oryzae*). *Jorn. Técn. Nac. del Cultivo de Arroz. Concordia, Entre Ríos*. Pp 119-127. *Resultados Experimentales 2012-2013*. Vol XXII. Ed. Fundación Proarroz. Pp 67-74. ISSN 1853-8754. 2013.
- [17] Piccoli, F.. *Computación de alto desempeño en GPU*. La Plata: EDULP, ISBN:978-950-34-0759-2. 2011.
- [18] Raja,R.; Nagwanshi,K.; Kumar, S.; Ramya Laxmi; K.. *Data Mining and Machine Learning Applications*. John Wiley & Sons. ISBN 9781119791782. 2022
- [19] Ross, T.. *Fuzzy logic with engineering applications*. John Wiley & Sons, ISBN 9781119235866. 2016.
- [20] Schab, E., Casanova, C., Piccoli, F.. Solving an Instance of a Routing Problem Through Reinforcement Learning and High Performance Computing. In: Rucci,E., Naiouf,M., Chichizola,F., De Giusti,L., De Giusti,A.(eds) *Cloud Computing, Big Data & Emerging Topics*. JCC-BD&ET 2022. *Comm. in Computer and Information Science*, vol 1634. Springer, Cham. https://doi.org/10.1007/978-3-031-14599-5_8. 2022.
- [21] Schab, Esteban, Lopresti, Mariela, Miranda, Natalia Carolina, Casanova, Carlos, Piccoli, María Fabiana. "Toma de decisiones en sistemas de eventos mediante inteligencia computacional y computación de alto desempeño". XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza). Red de Universidades con Carreras en Informática. ISBN: 978-987-48222-3-9. Pp 535-539. Universidad Champagnat, Mendoza. Abril 2022.
- [22] SIBER. Inf. Sup. Sembrada y Prod. del Arroz. Camp. 2020/21. <https://www.bolsacer.org.ar/Fuentes/siberd.php?Id=1259>. 2021.
- [23] Siddique, N.; Adeli, H.. *Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*. John Wiley and Sons. ISBN 9781118534816. 2013.
- [24] Skoda, P.; Fathalrahman, A. *Knowledge discovery in big data from astronomy and Earth observation : astrogeoinformatics*. Elsevier. ISBN 9780128191545. 2020.
- [25] Wu, K., Simon, H.D.. *High-Performance Computational Intelligence and Forecasting Technologies*. Lawrence Berkeley National Laboratory. Ch: 22. Wiley. ISBN: 9781119482086. 2021.
- [26] Yoshino R. Ecological studies on the penetration of rice blast fungus, *Pyricularia oryzae* into leaf epidermal cells. *Bulletin of the Hokuriku National Agricultural Experiment Station*. ;22: 163–221. 1979
- [27] Zadeh, L. A.. "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 1, pp. 28-44, 1973.
- [28] Zimmermann, H. . Testability and meaning of mathematical models in social sciences. *Mathematical Modelling* 1(1). ISSN 0270-0255. Pp 123-139. 1980.