
**MODELO PREDICTIVO DE LOS RASGOS FENOTÍPICOS DEL
TRIGO: UNA METODOLOGÍA SISTEMÁTICA**

**MARIANELA ANDREA ITURRIAGA JIMÉNEZ
INGENIERO CIVIL EN COMPUTACIÓN**

RESUMEN

El cambio climático y la creciente población mundial plantean la necesidad de adaptar cultivos a diferentes condiciones. Hoy en día, los científicos y las empresas se enfrentan a un reto importante en el aumento de la productividad del suelo con el fin de obtener más alimentos. Uno de los cereales más utilizados es el trigo debido a sus propiedades nutricionales deseables. Científicos en Chile están llevando a cabo experimentos con diferentes variaciones del gen del trigo para encontrar uno que sea más rico y nutritivo. El proceso tradicional consiste en cultivar diferentes tipos de trigo y observar a través del tiempo su rendimiento. Sin embargo, esta tarea es costosa, ya que requiere mano de obra especializada y equipos costosos. El objetivo de esta investigación es estimar analíticamente el rendimiento del trigo sobre la base de información estadística. Esto se logra construyendo un modelo de regresión que sea capaz de predecir la capacidad de producción basada en mediciones de la reflectancia espectral de los individuos. Diseñamos un modelo de regresión basado en datos de reflectancia espectral para predecir el rendimiento del trigo. Como parte de la metodología, procesamos previamente los datos, eliminando las instancias con datos incompletos, encontrando valores atípicos, reduciendo la dimensionalidad de los datos y construyendo los regresores. Esta metodología se explica de la manera más genérica posible, con la esperanza de que sea un recurso útil para futuros investigadores que quieran reproducir nuestro método para estimar el rendimiento del trigo o cualquier otro tipo de planta.

ABSTRACT

Climate change and the increasing world population bring about the need of adapting crops to different conditions. Nowadays, scientists and companies face a major challenge in increasing the productivity of the soil so as to obtain more food. One of the most widely used cereals is wheat, due to its desirable nutrition properties. Scientist in Chile are carrying out experiments using different variation of the wheat gene, in order to find a strain that is wealthier and more nutritious. The traditional process consists in growing different types of wheat and observe its performance over a lengthy time span. This task is costly, since it necessitates specialized labor and expensive equipment. The goal of this research is to analytically estimate the performance of wheat based on statistical information. This is achieved by constructing a regression model that is able to predict the production capacity based on measurements of the spectral reflectance of the individuals. We propose a regression model based on spectral reflectance data which allows to predict the performance of the wheat. As part of the methodology, we pre-process the data, removing incomplete data, finding outliers, reducing the dimensionality of the data, and building the regressors. This methodology is explained as generic as possible, in hopes that it will be used as a useful resource for future researchers that want to reproduce our method for estimating the performance of wheat or any other type of plant.