Chapter 24

Temporal Data Analysis and Mining Methods for Modelling the Climate Change Effects on Malaysia's Oil Palm Yield at Different Regional Scales

Subana Shanmuganathan
Auckland University of Technology, New Zealand

Ajit Narayanan
Auckland University of Technology, New Zealand

Nishantha Priyanka Kumara Medagoda
Auckland University of Technology, New Zealand

ABSTRACT

Space and time related data generated is becoming ever more voluminous, noisy and heterogeneous outpacing the research efforts in the domain of climate. Nevertheless, this data portrays recent climate/weather change patterns. Thus, insightful approaches are required to overcome the challenges when handling the so called "big data" to unravel the recent unprecedented climate change in particular, its variability, frequency and effects on key crops. Contemporary climate-crop models developed at least two decades ago are found to be unsuitable for analysing complex climate/weather data retrospectively. In this context, the chapter looks at the use of scalable time series analysis, namely ARIMA (Autoregressive integrated moving average) models and data mining techniques to extract new knowledge on the climate change effects on Malaysia's oil palm yield at the regional and administrative divisional scales. The results reveal recent trends and patterns in climate change and its effects on oil palm yield impossible otherwise e.g. Traditional statistical methods alone.

DOI: 10.4018/978-1-5225-0803-8.ch024

INTRODUCTION

Research into addressing the major issues relating to the nature, storage, pre-processing, analysis and interpretation of big data generated in the climate domain is severely lagging behind the current data growth rate (Faghmous, 2014). Indeed, there is "a stark contrast from other fields such as advertising and electronic commerce where big data has been a great success story" (Faghmous & Kumar, 2014 p1). Despite this lack of research in this domain, space and time related information continues to grow at unprecedented rates, so are the noise and heterogeneity in the data acquired by a multitude of telemetry in-situ instruments and sensors. With the introduction of more sophisticated data acquisition devices and equipment, such as new satellites and radars, climate and weather data holdings (depositories/ centres) are anticipated to grow exponentially. For instance, from 2004 to 2013, digital archives of National Climatic Data Center (NCDC), the world's largest weather/ climate data archives, grew from 2 to14 petabytes (PB[1]) (NCDC, 2014). Such massive volumes of noisy and heterogeneous data increasingly present more complex constraints when handling and analysing the data especially, using conventional methodologies. The recently acquired climate/weather data in addition to increased volume related issues, also encompasses auto- and cross-correlations between input variables. More specifically, the data presents severe limitations with the existing learning algorithms that make implicit or explicit independence assumptions about the input data.

Besides the lag in research relating to climate data processing and analysis, contemporary climate-crop models are described to be outdated (Rotter, Carter, Olesen, & Porter, 2011). Conventional methods currently in use for modelling climate-crop interactions were developed at least two decades ago. These methods originally designed for prospective analysis of survey data reflect "reductionism[2]". They do not address the major impediments relating to retrospective analysis of "big data" that depict the recent variability in weather and climatic conditions. For instance, the frequency and extremity embedded in the raw data in the form of trends, patterns and correlations cannot be analysed using conventional modelling approaches or techniques (Callebaut, 2012). Hence, *digital ecosystem* approaches (the application of holistic[3]/systems concepts in computing) are required to analyse multi-sourced data sets collectively/ integrated with other data (Mattmann, 2013), in this case with crop yield. This is seen as vital to unravel the recent change patterns in climate and weather to model/predict the change effects on key crops for decision/policy making at many critical levels and scales i.e., from daily farm management operations to policy making relating to the world food security now and in the near future (UCAR Office of the President Developments and Partnerships, 2014).

Yet another aspect that makes climate-crop modelling more complicated is the fact that climate change effects are observed to be somewhat stochastic depending on the climate-crop regime (Köstner, et al., 2014). The effects could be positive or negative, negligible or extensive depending on the change in the regional climate, to be precise, on the local weather, site quality, and land use and management practices adopted. Hence, any practical measure to adapt to the local weather change by way of plant management practices i.e., to irrigate rain fed crops, extensive knowledge on potential and specific impacts of climate change at the local site, regional, as well as global scales are required. In addition to these factors, different crops respond very differently to the same climate/ weather change, in the case of oil palm, frond emission, flowering and fruit production are the stages that get affected. The stages can also be affected by internal factors i.e., plant age, tolerance to drought, simultaneously reflecting seasonal maxima/ minima. Finally, diseases and pests play a major role in the oil palm production and their effects on oil palm under changing climate too are anticipated to be unprecedented (Paterson, Sariah, & Lima, 2013).