

Shoshi, Humayra; Hanson, Erik; Nganje, William; SenGupta, Indranil

Article

Stochastic analysis and neural network-based yield prediction with precision agriculture

Journal of Risk and Financial Management

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Shoshi, Humayra; Hanson, Erik; Nganje, William; SenGupta, Indranil (2021) : Stochastic analysis and neural network-based yield prediction with precision agriculture, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 14, Iss. 9, pp. 1-17, <https://doi.org/10.3390/jrfm14090397>

This Version is available at:

<https://hdl.handle.net/10419/258501>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Article

Stochastic Analysis and Neural Network-Based Yield Prediction with Precision Agriculture

Humayra Shoshi ¹, Erik Hanson ², William Nganje ² and Indranil SenGupta ^{1,*}

¹ Department of Mathematics, North Dakota State University, Fargo, ND 58108-6050, USA; humayra.shoshi@ndsu.edu

² Department of Agribusiness and Applied Economics, North Dakota State University, Fargo, ND 58108-6050, USA; erik.drevlow.hanson@ndsu.edu (E.H.); william.nganje@ndsu.edu (W.N.)

* Correspondence: indranil.sengupta@ndsu.edu

Abstract: In this paper, we propose a general mathematical model for analyzing yield data. The data analyzed in this paper come from a characteristic corn field in the upper midwestern United States. We derive expressions for statistical moments from the underlying stochastic model. Consequently, we illustrate how a particular feature variable contributes to the statistical moments (and in effect, the characteristic function) of the target variable (i.e., yield). We also analyze the data with neural network techniques and provide two methods of data analysis. This mathematical model and neural network-based data analysis allow for better understanding of the variability within the data set, which is useful to farm managers attempting to make current and future decisions using the yield data. Lenders and risk management consultants may benefit from the insights of this mathematical model and neural network-based data analysis regarding yield expectations.

Keywords: neural networks; precision agriculture; statistical moments; yield; categorical data



Citation: Shoshi, Humayra, Erik Hanson, William Nganje, and Indranil SenGupta. 2021. Stochastic Analysis and Neural Network-Based Yield Prediction with Precision Agriculture. *Journal of Risk and Financial Management* 14: 397. <https://doi.org/10.3390/jrfm14090397>

Academic Editor: Zbigniew Palmowski

Received: 4 August 2021

Accepted: 20 August 2021

Published: 25 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The International Society of Precision Agriculture adopted the following definition of precision agriculture in 2019 (see [The International Society of Precision Agriculture \(n.d.\)](#)): “Precision Agriculture is a management strategy that gathers, processes, and analyzes temporal, spatial, and individual data and combines it with other information to support management decisions according to estimated variability for improved resource using efficiency, productivity, quality, profitability, and sustainability of agricultural production”. Precision agriculture has emerged as a central tool to address current challenges in agricultural sustainability and profitability. Various methodologies of data-science, such as machine learning, have been implemented with this cutting edge technology. With the use of artificial intelligence and data-science, farmer managers can get the precise data that convey all the information related to the optimum health and productivity of the crops, thereby enabling informed decision-making.

There is an emerging literature on applications of data-science to agriculture. In [Lemley et al. \(2017\)](#), machine learning and deep learning techniques are implemented to solve both precision agriculture related problems and build better, smarter consumer devices and services. In [Sharma et al. \(2021\)](#), a systematic review of various machine learning applications in the field of agriculture is provided. The prediction of soil parameters such as organic carbon and moisture content, crop yield prediction, disease and weed detection in crops and species detection is presented. In addition, this paper demonstrates how knowledge-based agriculture can improve the sustainable productivity and quality of the product. In [Bauer et al. \(2019\)](#), the authors present an automated and open-source analytic platform that combines computer vision, and modular software engineering in order to measure yield-related phenotypes from ultra-large aerial imagery. An analysis is developed to map lettuce size distribution across the field, based on which associated global

positioning system (GPS) tagged harvest regions have been identified to enable growers and farmers to conduct precision agricultural practices. In [Chlingaryan et al. \(2018\)](#), the authors analyze research developments performed within the last 15 years on machine learning based techniques for accurate crop yield prediction and nitrogen status estimation. In [Trebooux and Genoud \(2020\)](#) the authors present the performances of machine learning algorithms on aerial images object detection for high-precision agriculture. The proposed approach in this paper improves the object detection and obtain an accuracy of 94.27%. In [Hornig et al. \(2020\)](#) the authors propose a study of harvesting system based on the Internet of Things technology and smart image recognition. The proposed model is implemented for crop detection by collecting and tagging images. It is shown that the proposed model training has a mean average precision of 84%, which is better than the other existing models. In [Bauer et al. \(2019\)](#) a common problem for apple orchards, namely, the attack of the codling moth, is studied. It is shown that a data-science based near sensor neural network algorithms can be implemented to automatically detect the codling moth. The performance of this system is evaluated, and power consumption ideas are discussed for achieving the zero energy balance of the system. Finally, in [Addey et al. \(2021\)](#) the authors examine the implications of risks, uncertainties and random events on the prediction of crop yields.

Motivated by all these studies, we propose a general mathematical model for analyzing yield data. It is shown that a special case of such a model can be a generalized version of the well-known Barndorff–Nielsen and Shephard model. If the statistical moments of the yield data can be obtained from the stochastic model, then it can play a crucial role in understanding the empirical data set. In fact, this will illustrate how a particular feature variable contributes to the statistical moments (and in effect, the characteristic function) of the target variable (i.e., yield). Consequently, we derive expressions for statistical moments from the underlying stochastic model. This model allows for better understanding of the variability within the data set, which is useful to farm managers attempting to make current and future decisions using the yield data. In addition, this model may help to identify errors in yield data collection and recording. Finally, lenders and risk management consultants may benefit from this model's insights regarding yield expectations.

The data analyzed in this paper come from a characteristic corn field in the upper midwestern United States. The data were collected in 2010. Each observation in the data set represents two seconds of harvesting as a combine harvester traveling through the field. Each observation is associated with a precise latitude and longitude where it was recorded in the field. Yld Vol(Dry)(bu/ac) is the quantity of corn harvested during the 2-seconds interval. It is the variable predicted by the model. The mean Yld Vol(Dry)(bu/ac) for the field is 161.06 and the median Yld Vol(Dry)(bu/ac) is 172.07. Yld Mass(Wet)(lb/ac), Yld Mass(Dry)(lb/ac), and Crop Flw(M)(lb/s) are also measures of corn yield during each 2-seconds interval, but these measures have not been converted to bushels. Furthermore, Yld Mass(Wet)(lb/ac) has not been corrected for moisture, which is represented by the crop's relative moisture, Moisture(%), and Crop Flw(M)(lb/s) measures corn harvested per second rather than per acre. During each 2-seconds interval, the combine harvested travels a variable distance, Distance(ft), which is determined by the Speed(mph) that the combine is traveling during the interval. The combine's speed and the width of the combine's header during the 2-seconds interval, Swth Wdth(ft), determine how many acres could be harvested in one hour at those rates (Prod(ac/h)). We implement neural network algorithms to analyze the dry yield volume based on the other observed and significant feature variables.

The organization of the paper is as follows. In Section 2, we propose a general statistical model, and analyze a special case of it—a generalized Barndorff–Nielsen and Shephard model. Theoretical results related to the statistical moments of the *target variable* are discussed in detail. In Section 3, we provide the description of the data. We analyze the data with neural network techniques and provide two methods of data analysis. Finally, a brief conclusion is provided in Section 4.

2. Mathematical Model and Analysis

2.1. General Framework

For the empirical data, we assume that there is one *target variable* and n *feature variables*. We model the *target variable* S_t by

$$S_t = S_0 e^{X_t}, \quad \text{where} \quad dX_t = b_t dt + \sum_{i=1}^n \theta_t^{(i)} \left(\sigma_t dW_t^{(i)} + dJ_t^{(i)} \right), \quad (1)$$

where b_t is a deterministic function of t , $W_t^{(i)}$, $i = 1, \dots, n$, are independent Brownian motions, and $J_t^{(i)}$ is the jump process with intensity λ_i , $i = 1, \dots, n$. We assume that $W_t^{(i)}$ and $J_t^{(i)}$, for $i = 1, \dots, n$, are independent. The coefficients $\theta_t^{(i)}$, at every t satisfy $\sum_{i=1}^n (\theta_t^{(i)})^2 = 1$. In addition to that, σ_t is assumed to be stochastic, and its dynamics are governed by

$$d\sigma_t^2 = F(\sigma_t^2, \beta_t^{(1)} H_t^{(1)}, \beta_t^{(2)} H_t^{(2)}, \dots, \beta_t^{(n)} H_t^{(n)}), \quad (2)$$

for an appropriate function F , where $H_t^{(j)}$, for $j = 1, \dots, n$, are jump processes with intensities μ_j , $j = 1, \dots, n$. The coefficients $\beta_t^{(j)}$, at every t satisfy $\sum_{j=1}^n (\beta_t^{(j)})^2 = 1$. For simplicity, for the rest of the proposal, we assume $\theta^{(i)} = \beta^{(i)}$, for $i = 1, \dots, n$.

For a special case of (1), we assume that the individual dynamics of a *feature variable* is given by $e^{Y_t^{(i)}}$, where $dY_t^{(i)} = \sigma_t dW_t^{(i)} + dJ_t^{(i)}$, $i = 1, \dots, n$. From (1), we obtain $dX_t = b_t dt + \sum_{i=1}^n \theta_t^{(i)} dY_t^{(i)}$. Hence, $\theta_t^{(i)}$ represents the ‘‘importance factor’’ for the i -th *feature variables*, for $i = 1, \dots, n$. We observe, that if $\sum_{i=1}^n (\theta_t^{(i)})^2 = 1$, then $\sum_{i=1}^n \theta_t^{(i)} dW_t^{(i)}$ can be represented by dB_t , where B_t is a Brownian motion. Consequently, (1) can be written as

$$S_t = S_0 e^{X_t}, \quad \text{where} \quad dX_t = b_t dt + \sigma_t dB_t + \sum_{i=1}^n \theta_t^{(i)} dJ_t^{(i)}. \quad (3)$$

The expression (3) provides an alternative explanation for the coefficients $\theta_t^{(i)}$, $i = 1, \dots, n$, and those will be computed in the numerical section. Those represent the significance in terms of big fluctuations (or ‘‘jumps’’) of the i -th ingredient *feature process* $Y_t^{(i)}$.

We write $J_t^{(i)}$ in terms of integrals with respect to Poisson random measures $N^{(i)}(dt, dx)$, for $i = 1, \dots, n$. Consequently,

$$J_t^{(i)} = \int_0^t \int_{\mathbb{R}} x N^{(i)}(dt, dx).$$

Hence, (3) can be written as

$$S_t = S_0 e^{X_t}, \quad \text{where} \quad dX_t = b_t dt + \sigma_t dB_t + \sum_{i=1}^n \theta_t^{(i)} \int_{\mathbb{R}} x N^{(i)}(dt, dx). \quad (4)$$

2.2. Special Case: A Generalized Barndorff–Nielsen & Shephard Model

There are some special cases of the proposed model that are studied in the literature in connection to the financial market; for example, the Barndorff–Nielsen and Shephard model (BN-S model). For such a model, the *target variable* is the stock, see, [Barndorff-Nielsen and Shephard \(2001a, 2001b\)](#); [Habtemicael and SenGupta \(2016\)](#); [Issaka and SenGupta \(2017\)](#) or the commodity price (see, [Roberts and SenGupta \(2020\)](#); [SenGupta et al. \(2019\)](#); [Shoshi and SenGupta \(2021\)](#); [Wilson et al. \(2019\)](#)), $S = (S_t)_{t \geq 0}$. On some filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{0 \leq t \leq T}, \mathbb{P})$ it is modeled by

$$S_t = S_0 \exp(X_t), \quad (5)$$

$$dX_t = (\mu + \beta \sigma_t^2) dt + \sigma_t dW_t + \rho dZ_{\lambda t}, \quad (6)$$

$$d\sigma_t^2 = -\lambda\sigma_t^2 dt + dZ_{\lambda t}, \quad \sigma_0^2 > 0, \tag{7}$$

where the parameters $\mu, \beta, \rho, \lambda \in \mathbb{R}$ with $\lambda > 0$ and $\rho \leq 0$ and r is the risk-free interest rate where a stock or commodity (the *target variable*) is traded up to a fixed horizon date T . In the above model, W_t is a Brownian motion, and the process $Z_{\lambda t}$ is a subordinator. Additionally, W and Z are assumed to be independent, and (\mathcal{G}_t) is assumed to be the usual augmentation of the filtration generated by the pair (W, Z) . We consider a special case of (4), where $dZ_s^{(i)} = \frac{1}{\rho} \int_0^\infty xN^{(i)}(ds, dx), i = 1, \dots, n$ are subordinators. Making a scaling in the time variable, we define $s = \lambda t$, for $\lambda > 0$. Then, we obtain $dZ_{\lambda t}^{(i)} = \frac{1}{\rho} \int_0^\infty xN^{(i)}(\lambda dt, dx), i = 1, \dots, n$, as subordinators. Consequently, we consider $S = (S_t)_{t \geq 0}$ on some risk-neutral probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{Q})$, given by (4). We consider the case that is more aligned with Nicolato and Venardos (2003). In this case, we assume that the generalized version of (6) takes the form

$$dX_t = (B - \frac{1}{2}\sigma_t^2) dt + \sigma_t dW_t + \rho \sum_{i=1}^n \theta_t^{(i)} dZ_{\lambda t}^{(i)}, \tag{8}$$

where $Z^{(i)}, i = 1, \dots, n$ are independent subordinators. For the drift term, comparing with (6), we thus have $\mu = B$ and $\beta = -\frac{1}{2}$. Additionally, we assume that (\mathcal{F}_t) is assumed to be the usual augmentation of the filtration generated by $(W, Z^{(i)}), i = 1, \dots, n$. In this case, (7) will be given by

$$d\sigma_t^2 = -\lambda\sigma_t^2 dt + \sum_{i=1}^n \theta_t^{(i)} dZ_{\lambda t}^{(i)}, \quad \sigma_0^2 > 0. \tag{9}$$

The solution of (9) can be explicitly written as

$$\sigma_t^2 = e^{-\lambda t} \sigma_0^2 + \int_0^t e^{-\lambda(t-s)} \sum_{i=1}^n \theta_s^{(i)} dZ_{\lambda s}^{(i)}. \tag{10}$$

The *integrated variance* over the time period $[t, T]$ is given by $\sigma_T^2 = \int_t^T \sigma_s^2 ds$, and a straight-forward calculation shows

$$\sigma_T^2 = \epsilon(t, T)\sigma_t^2 + \int_t^T \epsilon(s, T) \sum_{i=1}^n \theta_s^{(i)} dZ_{\lambda s}^{(i)}, \tag{11}$$

where

$$\epsilon(s, T) = (1 - \exp(-\lambda(T - s)))/\lambda, \quad t \leq s \leq T. \tag{12}$$

We derive a general expression for the characteristic function of the conditional distribution of the log-asset price process appearing in the stochastic model given by Equations (5), (8) and (9).

We quote the following result from Nicolato and Venardos (2003), which is known as the “key formula”. This result will play a significant role in the next section.

Lemma 1. *Let Z be a subordinator with cumulant transform κ , and let $f : \mathbb{R}_+ \rightarrow \mathbb{C}$ be a complex-valued, left-continuous function such that $Re(f) \leq 0$. Then*

$$E \left[\exp \left(\int_0^t f(s) dZ_{\lambda s} \right) \right] = \exp \left(\lambda \int_0^t \kappa(f(s)) ds \right).$$

We also note that

$$\int_0^t \sigma_s^2 ds = \frac{\sigma_0^2}{\lambda}(1 - e^{-\lambda t}) + \frac{1}{\lambda} \sum_{i=1}^n \int_0^t (1 - e^{-\lambda(t-s)}) \theta_s^{(i)} dZ_{\lambda s}^{(i)} \\ \stackrel{d}{=} \frac{\sigma_0^2}{\lambda}(1 - e^{-\lambda t}) + \frac{1}{\lambda} \sum_{i=1}^n \int_0^{\lambda t} (\theta_s^{(i)} - \theta_{t-\frac{s}{\lambda}}^{(i)} e^{-s}) dZ_s^{(i)}, \tag{13}$$

where “ $\stackrel{d}{=}$ ” represents equality in distribution.

Statistical moments are useful quantities, which when computable, can give significant insight into the underlying empirical data. Consequently, it is useful to obtain analytical expressions for statistical moments for the underlying model. In connection to the BN-S model, statistical moments for S_t are derived in Ihsan and SenGupta (2018). With this motivation, we derive a couple of results related to certain moments of the target variable, S_t . The first result is simple but restrictive. This result is subsequently generalized. For the rest of the paper, we take $N \in \mathbb{N}$, where \mathbb{N} is the set of natural numbers.

Theorem 1. For a given $\lambda > 0$ and $\rho < 0$, if N satisfies

$$N \leq 1 - 2\lambda\rho, \tag{14}$$

then the N -th moment of S_t with respect to the measure \mathbb{Q} is given by

$$E^{\mathbb{Q}}[e^{NX_t}] = a(t, N) \prod_{i=1}^n \exp \left[\lambda \int_0^t \kappa^{(i)} \left(\theta_s^{(i)} \left(N\rho + \frac{N(N-1)}{2\lambda} (1 - e^{-\lambda(t-s)}) \right) \right) ds \right], \tag{15}$$

where

$$a(t, N) = e^{NBt + \frac{N(N-1)}{2\lambda}(1 - e^{-\lambda t})\sigma_0^2}, \tag{16}$$

and $\kappa^{(i)}(\cdot)$ is the cumulant transform of $Z_1^{(i)}$, $i = 1, \dots, n$, with respect to \mathbb{Q} .

Proof. Let \mathcal{G} denote the σ -algebra generated by the Background Driving Lévy Process, BDLP, Z up to time t . We observe

$$E^{\mathbb{Q}}[e^{NX_t}] = E^{\mathbb{Q}}[e^{N(Bt - \frac{1}{2} \int_0^t \sigma_u^2 du + \int_0^t \sigma_u dW_u + \rho \sum_{i=1}^n \int_0^t \theta_u^{(i)} dZ_{\lambda u}^{(i)})}] \\ = e^{NBt} E^{\mathbb{Q}} \left[e^{\frac{N^2 - N}{2} \int_0^t \sigma_u^2 du + N\rho \sum_{i=1}^n \int_0^t \theta_u^{(i)} dZ_{\lambda u}^{(i)}} \right].$$

Using the first equality in (13) for the integrated variance process, we obtain

$$E^{\mathbb{Q}}[e^{NX_t}] = e^{NBt + \frac{N(N-1)}{2\lambda}(1 - e^{-\lambda t})\sigma_0^2} E^{\mathbb{Q}} \left[e^{\sum_{i=1}^n \int_0^t \theta_s^{(i)} \left(N\rho + \frac{N(N-1)}{2\lambda} (1 - e^{-\lambda(t-s)}) \right) dZ_{\lambda s}^{(i)}} \right] \\ = e^{NBt + \frac{N(N-1)}{2\lambda}(1 - e^{-\lambda t})\sigma_0^2} \prod_{i=1}^n E^{\mathbb{Q}} \left[e^{\int_0^t \theta_s^{(i)} \left(N\rho + \frac{N(N-1)}{2\lambda} (1 - e^{-\lambda(t-s)}) \right) dZ_{\lambda s}^{(i)}} \right], \tag{17}$$

where in the last step, we use the independence of $Z^{(i)}$, $i = 1, \dots, n$. If N satisfies $\left(N\rho + \frac{N(N-1)}{2\lambda} \right) \leq 0$, that is, N satisfies (14), then for $0 \leq s \leq t$,

$$\left(N\rho + \frac{N(N-1)}{2\lambda} (1 - e^{-\lambda(t-s)}) \right) \leq 0.$$

Consequently, by using Lemma 1, we obtain (15). \square

Next, we derive a result without any restriction on N . For this paper, we denote the random measure associated with the jumps of a process A_t , and the Lévy density of A_t , by $J_A(\cdot, \cdot)$ and $\nu_A(\cdot)$, respectively. The compensator for $J_A(dt, dx)$ is $\nu_A(dx) dt$, and we define

$\tilde{J}_A(dt, dx) = J_A(dt, dx) - \nu_A(dx) dt$. For the proof of Theorem 3, we use the following version of the Girsanov theorem. The proof may be found in Øksendal and Sulem (2007).

Theorem 2. Let $u(t)$ and $\theta(t, z) \leq 1$ be predictable processes such that the process

$$Z(t) := \exp \left(- \int_0^t u(s) dW_s - \frac{1}{2} \int_0^t u^2(s) ds + \int_0^t \int_{\mathbb{R}} \ln(1 - \theta(s, z)) \tilde{J}_Z(ds, dz) + \int_0^t \int_{\mathbb{R}} [\ln(1 - \theta(s, z)) + \theta(s, z)] \nu_Z(dz) ds \right),$$

exists for $0 \leq t \leq T$ and satisfies $E^{\mathbb{P}_1}[Z(T)] = 1$. Define the probability measure \mathbb{P}_2 by $d\mathbb{P}_2 = Z(T)d\mathbb{P}_1$. Then $u(t)dt + dW_t$ is a Brownian motion, and $\theta(t, z)\nu(dz)dt + \tilde{N}(dt, dz)$ is a compensated Poisson random process with respect to \mathbb{P}_2 .

We now proceed to prove a more general result related to the moments of S_t , with respect to the BN-S model. This result does not assume any restriction on $N \in \mathbb{N}$.

For the short-term, we assume $\theta_s^{(i)}$ to be a positive constant and $\theta_s^{(i)} = \theta^{(i)} > 0$, $i = 1, \dots, n$.

Theorem 3.

$$E^{\mathbb{Q}}(S_t^N) = \tau_2(t, N) \prod_{i=1}^n \exp \left(\int_0^{\lambda t} \kappa^{(i)} \left(-\theta^{(i)} \frac{N(N-1)}{2\lambda} e^{-s} \right) ds \right), \tag{18}$$

where $\tau_2(t, N)$ is a deterministic function of t , and $\kappa^{(i)}(\cdot)$ is the cumulant generating function for $Z_1^{(i)}$, with respect to \mathbb{Q} , where $i = 1, \dots, n$.

We remark that an explicit form of a particular case of the function $\tau_2(t, N)$ in (18) will be found in the Corollary 4.

Proof. Let U be a subordinator with Poisson measure $J_U(\cdot, \cdot)$. We will characterize the process later in this proof. Consider the stochastic differential equation

$$\frac{dM_t}{M_t} = \alpha dt + \beta dW_t + \int_{\mathbb{R}_+} \gamma(y) \tilde{J}_U(\lambda dt, dy), \quad M_0 = 1,$$

where α, β are constants and $\gamma(y) > -1$. Consequently, (see Øksendal and Sulem (2007)) we obtain

$$M_t = \exp \left[\left(\alpha - \frac{1}{2}\beta^2 \right) t + \beta W_t + \int_0^{\lambda t} \int_{0 < y < 1} (\ln(1 + \gamma(y)) - \gamma(y)) \nu_U(dy) ds + \int_0^{\lambda t} \int_{\mathbb{R}_+} \ln(1 + \gamma(y)) \tilde{J}_U(ds, dy) \right]. \tag{19}$$

Set $\beta = 0$ in (19) to obtain

$$M_t = \exp \left[\alpha' t + \int_0^{\lambda t} \int_{\mathbb{R}_+} (\ln(1 + \gamma(y)) - \gamma(y)) \nu_U(dy) ds + \int_0^{\lambda t} \int_{\mathbb{R}_+} \ln(1 + \gamma(y)) \tilde{J}_U(ds, dy) \right],$$

where $\alpha' = \alpha - \lambda \int_{y \geq 1} (\ln(1 + \gamma(y)) - \gamma(y)) \nu_U(dy)$. We choose α in such a way that

$$\alpha' = 0.$$

Then by Cont and Tankov (2004) (Proposition 8.23), we obtain M_t as a martingale. Consider a new measure $d\mathbb{T}(t) = M_t d\mathbb{Q}(t)$. Note that with respect to \mathbb{T} , the Brownian motion W_t still remains the same.

Then

$$\begin{aligned}
 E^{\mathbb{Q}}(S_t^N) &= E^{\mathbb{T}}\left[\frac{1}{M_t}S_t^N\right] \\
 &= E^{\mathbb{T}}\left[\exp\left(-\int_0^{\lambda t}\int_{\mathbb{R}_+}(\ln(1+\gamma(y))-\gamma(y))\nu_U(dy)ds-\int_0^{\lambda t}\int_{\mathbb{R}_+}\ln(1+\gamma(y))\tilde{J}_U(ds,dy)\right.\right. \\
 &\quad \left.\left.+Nbt-\frac{N}{2}\int_0^t\sigma_s^2ds+N\int_0^t\sigma_s dW_s+\rho N\sum_{i=1}^n\theta^{(i)}\int_0^{\lambda t}\int_{\mathbb{R}_+}yJ_Z^{(i)}(ds,dy)\right)\right] \\
 &= \tilde{\tau}(t,N)E^{\mathbb{T}}\left[\exp\left(-\int_0^{\lambda t}\int_{\mathbb{R}_+}\ln(1+\gamma(y))\tilde{J}_U(ds,dy)-\frac{N}{2}\int_0^t\sigma_s^2ds+N\int_0^t\sigma_s dW_s\right.\right. \\
 &\quad \left.\left.+\rho N\sum_{i=1}^n\theta^{(i)}\int_0^{\lambda t}\int_{\mathbb{R}_+}yJ_Z^{(i)}(ds,dy)\right)\right],
 \end{aligned}$$

where $\tilde{\tau}(t,N) = \exp\left(Nbt - \lambda t \int_{\mathbb{R}_+}(\ln(1+\gamma(y))-\gamma(y))\nu_U(dy)\right)$, is a deterministic function of t . Using $\tilde{J}_U(ds,dy) = J_U(ds,dy) - \nu_U(dy)ds$, we obtain

$$\begin{aligned}
 E^{\mathbb{Q}}(S_t^N) &= E^{\mathbb{T}}\left[\frac{1}{M_t}S_t^N\right] \\
 &= \tau(t,N)E^{\mathbb{T}}\left[\exp\left(-\int_0^{\lambda t}\int_{\mathbb{R}_+}\ln(1+\gamma(y))J_U(ds,dy)-\frac{N}{2}\int_0^t\sigma_s^2ds+N\int_0^t\sigma_s dW_s\right.\right. \\
 &\quad \left.\left.+\rho N\sum_{i=1}^n\theta^{(i)}\int_0^{\lambda t}\int_{\mathbb{R}_+}yJ_Z^{(i)}(ds,dy)\right)\right],
 \end{aligned}$$

where

$$\begin{aligned}
 \tau(t,N) &= \exp\left(Nbt - \lambda t \int_{\mathbb{R}_+}(\ln(1+\gamma(y))-\gamma(y))\nu_U(dy) + \lambda t \int_{\mathbb{R}_+}\ln(1+\gamma(y))\nu_U(dy)\right) \\
 &= \exp\left(Nbt + \lambda t \int_{\mathbb{R}_+}\gamma(y)\nu_U(dy)\right),
 \end{aligned}$$

is a deterministic function of t . We choose $-1 < \gamma(y)$, such that the following holds:

$$-\int_{\mathbb{R}_+}\ln(1+\gamma(y))J_U(ds,dy) + \left(\rho N + \frac{N(N-1)}{2\lambda}\right)\sum_{i=1}^n\theta^{(i)}\int_0^{\lambda t}\int_{\mathbb{R}_+}yJ_Z^{(i)}(ds,dy) = 0. \tag{20}$$

Consequently,

$$\begin{aligned}
 E^{\mathbb{Q}}(S_t^N) &= \tau(t,N)E^{\mathbb{T}}\left[\exp\left(-\frac{N}{2}\int_0^t\sigma_s^2ds+N\int_0^t\sigma_s dW_s-\frac{N(N-1)}{2\lambda}\sum_{i=1}^n\theta^{(i)}\int_0^{\lambda t}\int_{\mathbb{R}_+}yJ_Z^{(i)}(ds,dy)\right)\right].
 \end{aligned}$$

As before, let \mathcal{G} denote the σ -algebra generated by the BDLP $Z^{(i)}, i = 1, \dots, n$, up to time t . Then we obtain

$$\begin{aligned}
 E^{\mathbb{Q}}(S_t^N) &= \tau(t, N) E^{\mathbb{T}} \left[e^{-\frac{N}{2} \int_0^t \sigma_s^2 ds - \frac{N(N-1)}{2\lambda} \sum_{i=1}^n \theta^{(i)} \int_0^{\lambda t} \int_{\mathbb{R}^+} y J_Z^{(i)}(ds, dy)} E^{\mathbb{T}} \left[e^{N \int_0^t \sigma_s dW_s} \middle| \mathcal{G} \right] \right] \\
 &= \tau(t, N) E^{\mathbb{T}} \left[e^{-\frac{N}{2} \int_0^t \sigma_s^2 ds - \frac{N(N-1)}{2\lambda} \sum_{i=1}^n \theta^{(i)} \int_0^{\lambda t} \int_{\mathbb{R}^+} y J_Z^{(i)}(ds, dy)} e^{\frac{N^2}{2} \int_0^t \sigma_s^2 ds} \right] \\
 &= \tau(t, N) E^{\mathbb{T}} \left[\exp \left(\frac{N(N-1)}{2} \int_0^t \sigma_s^2 ds - \frac{N(N-1)}{2\lambda} \sum_{i=1}^n \theta^{(i)} \int_0^{\lambda t} \int_{\mathbb{R}^+} y J_Z^{(i)}(ds, dy) \right) \right] \\
 &= \tau(t, N) E^{\mathbb{T}} \left[\exp \left(\frac{N(N-1)}{2\lambda} (1 - e^{-\lambda t}) \sigma_0^2 + \frac{N(N-1)}{2\lambda} \sum_{i=1}^n \int_0^t (-\theta^{(i)} e^{-\lambda(t-s)}) dZ_{\lambda s}^{(i)} \right) \right] \\
 &= \tau_1(t, N) E^{\mathbb{T}} \left[\exp \left(\frac{N(N-1)}{2\lambda} \sum_{i=1}^n \int_0^t \int_{\mathbb{R}^+} (-\theta^{(i)} e^{-\lambda(t-s)}) y \tilde{J}_Z^{(i)}(\lambda ds, dy) \right) \right],
 \end{aligned}$$

where

$$\tau_1(t, N) = \tau(t, N) \exp \left(\frac{N(N-1)}{2\lambda} (1 - e^{-\lambda t}) \sigma_0^2 + \frac{N(N-1)}{2} \sum_{i=1}^n \theta^{(i)} \int_0^t \int_{\mathbb{R}^+} (-e^{-\lambda(t-s)}) y \nu_Z^{(i)}(dy) ds \right).$$

With respect to \mathbb{T} , using Girsanov’s theorem, we find that the compensated subordinator is (see Øksendal and Sulem (2007) (Theorem 1.35)) given by

$$\tilde{J}_{\mathbb{T}}^{(i)}(\lambda ds, dy) = -\lambda \gamma(y) \nu_Z^{(i)}(dy) ds + \tilde{J}_Z^{(i)}(\lambda ds, dy), \quad i = 1, \dots, n. \tag{21}$$

Consequently,

$$\begin{aligned}
 E^{\mathbb{Q}}(S_t^N) &= \tau_2(t, N) E^{\mathbb{T}} \left[\exp \left(\frac{N(N-1)}{2\lambda} \sum_{i=1}^n \int_0^t \int_{\mathbb{R}^+} (-\theta^{(i)} e^{-\lambda(t-s)}) y \tilde{J}_{\mathbb{T}}^{(i)}(\lambda ds, dy) \right) \right] \\
 &= \tau_2(t, N) E^{\mathbb{T}} \left[\exp \left(\frac{N(N-1)}{2\lambda} \sum_{i=1}^n \int_0^t (-\theta^{(i)} e^{-\lambda(t-s)}) dZ_{\lambda s}^{(i)} \right) \right] \\
 &= \tau_2(t, N) E^{\mathbb{T}} \left[\exp \left(\frac{N(N-1)}{2\lambda} \sum_{i=1}^n \theta^{(i)} \int_0^{\lambda t} (-e^{-s}) dZ_s^{(i)} \right) \right],
 \end{aligned}$$

where Z_t , in the final steps, is a subordinator with respect to \mathbb{T} (consequently, (13) is used), and

$$\tau_2(t, N) = \tau_1(t, N) \exp \left(\frac{N(N-1)}{2\lambda} \sum_{i=1}^n \theta^{(i)} \left(\int_0^t \int_{\mathbb{R}^+} (-e^{-\lambda(t-s)}) \lambda y (\gamma(y) \nu_Z^{(i)}(dy) - \nu_{\mathbb{T}}^{(i)}(dy)) ds \right) \right).$$

Thus, with the application of Lemma 1, we obtain

$$E^{\mathbb{Q}}(S_t^N) = \tau_2(t, N) \prod_{i=1}^n \exp \left(\int_0^{\lambda t} \kappa^{(i)} \left(-\theta^{(i)} \frac{N(N-1)}{2\lambda} e^{-s} \right) ds \right).$$

□

Next, we present an immediate corollary that follows from Theorems 1 and 3, by setting $N = 1$. Note that when $N = 1$, since $\rho < 0$, the condition (14) is satisfied.

Corollary 4.

$$\tau_2(t, 1) = \exp \left[Bt + \lambda \sum_{i=1}^n \kappa^{(i)} \left(\theta^{(i)} \rho \right) t \right],$$

where $\kappa^{(i)}(\cdot)$ is the cumulant generating function for $Z_1^{(i)}$, with respect to \mathbb{Q} , where $i = 1, \dots, n$.

3. Data Description and Analysis

The data analyzed in this paper come from a North Dakota corn field in 2010. For the cleaned-up data set, we have six significant *feature variables* and one *target variable*. The *feature variables* are:

1. Swth Wdth(ft): The width of the header in ft;
2. Distance(ft): The distance (in ft) travelled between two data points;
3. Crop Flw(M)(lb/s): The crop flow harvested per second between data points;
4. Moisture(%): Crop moisture in %;
5. Speed(mph): The speed of the combine;
6. Prod(ac/h): Combine productivity per hour.

The *target variable* is:

- Yld Vol(Dry)(bu/ac): Yield volume dry.

In other words, for the above *feature variables*, in (4), for simplicity we assign $\theta_t^{(i)}$ as independent of time, and $\theta_t^{(i)} = 1$, for $i = 1, \dots, 6$, and S_t represents the “Yld Vol(Dry)(bu/ac)”. Figures 1–5 provide various exploratory data analyses.

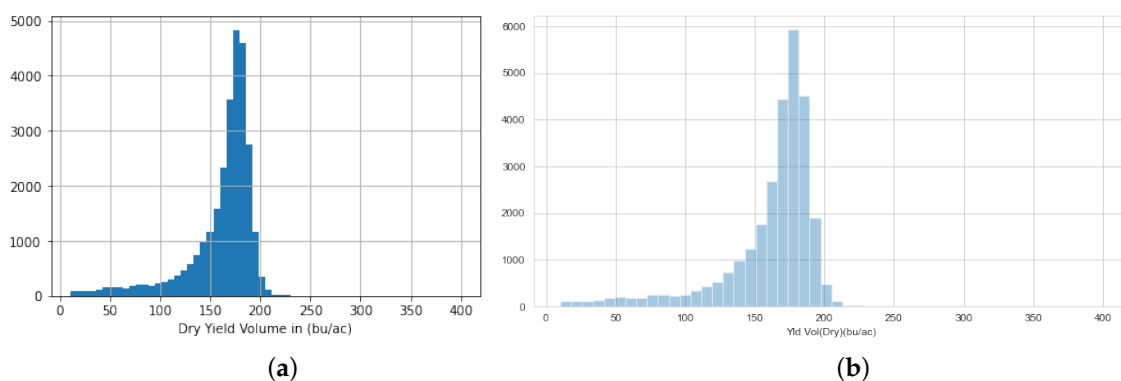


Figure 1. (a) Histogram for Yield Vol (Dry). (b) Distribution plot for Yield Vol (Dry).

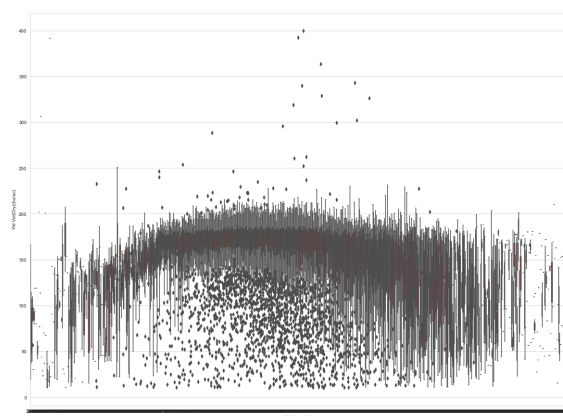


Figure 2. Box plot for Moisture (%) vs. Yld Vol (Dry).

For Figure 1, from the histogram and the distribution plot for Yield Vol (Dry), respectively, we observe that the distribution of the sample data points of feature Yld Vol (Dry) is left-skewed. It can also be noted that the maximum count of observation falls within the range of 150–200 of Yld Vol (Dry) (bu/ac) for both these plots. In Figures 2 and 3, for our exploratory data analysis, we implement box plots to visualize the measures of dispersion of our data set for the variable Yld Vol (Dry) with respect to two variables—Moisture (%) and Distance(ft), respectively. We observe from both plots that for a certain fixed interval, there is a cluster of outliers underneath the lower whisker of the plots. The inter-quartile

range tends to show an increasing pattern either right before or right after the big cluster of outliers. To help us visualize the distribution of a single variable and relationships between two variables, in Figure 4 we use the pairs plot. For our data set, the pairs plot in the diagonal gives us the univariate distribution via a histogram, and the scatter plots shows the bivariate relationship among variables. For Figure 5, the correlation matrix heatmap for our data set provides an overview on the relation among different variables (features). From the matrix heatmap color codes or annotated values, we can observe that the correlation coefficient between most features is low.

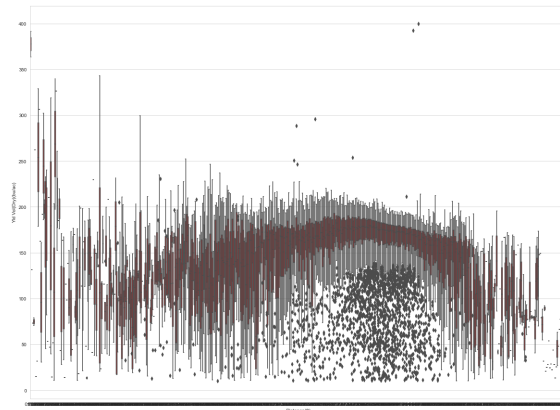


Figure 3. Box plot for Distance (ft) vs. Yld Vol (Dry).

The goal of the analysis is to predict the *target variable* (Yld Vol(Dry)(bu/ac)) based on the *feature variables*. We conduct the analysis in two different approaches.

Method 1: For this case, at first, we observe that for the “Yld Vol(Dry)(bu/ac)” variable, the maximum and minimum are 399.55 and 10.28, respectively. Consequently, we divide the target into eight equally spaced intervals, as [10, 60, 110, 160, 210, 260, 310, 360, 410].

Corresponding to each target variable, we create a list of three categorical data $[a, b, c]$, where $a, b, c \in \{0, 1\}$ by the following rule: if the target variable x is such that $10 \leq x < 60$, then $[a, b, c] = [0, 0, 0]$, if $60 \leq x < 110$, then $[a, b, c] = [0, 0, 1]$, if $110 \leq x < 160$, then $[a, b, c] = [0, 1, 0]$, if $160 \leq x < 210$, then $[a, b, c] = [0, 1, 1]$, if $210 \leq x < 260$, then $[a, b, c] = [1, 0, 0]$, if $260 \leq x < 310$, then $[a, b, c] = [1, 0, 1]$, if $310 \leq x < 360$, then $[a, b, c] = [1, 1, 0]$, and if $360 \leq x < 410$, then $[a, b, c] = [1, 1, 1]$.

After this, we train the neural network model with 80% data and test it on the remaining 20% data. We create a four-layered deep model with 42 nodes in the first layer, 30 nodes in the second layer, and 20 and 10 nodes in the last two layers, respectively. We use the activation function Rectified Linear Unit (ReLU) for the first input layer followed by the tanh activation function for the three hidden layers. For the output, we use the softmax activation function. To determine which hyper-parameter combination is most efficient for our model, we first conduct exploratory data analysis (EDA) to visualize the validation loss and accuracy of our model and then choose to train the model on 120 epochs, with a batch size of 32.

For the testing data, *true positive*, *true negative*, *false positive*, and *false negative* are denoted as TP, TN, FP, and FN, respectively. The following measurements are standard:

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

The *f1-score* gives the harmonic mean of precision and recall. The scores corresponding to every class gives the accuracy of the classifier in classifying the data points in that particular class compared to all other classes. The *support* is the number of samples of the true response that lie in that class. Table 1 provides the classification report for Method 1.

Table 1. Classification report.

Label	Precision	Recall	f1-Score	Support
a	0.53	0.99	0.69	99
b	0.99	1.00	0.99	5115
c	0.98	0.75	0.85	438

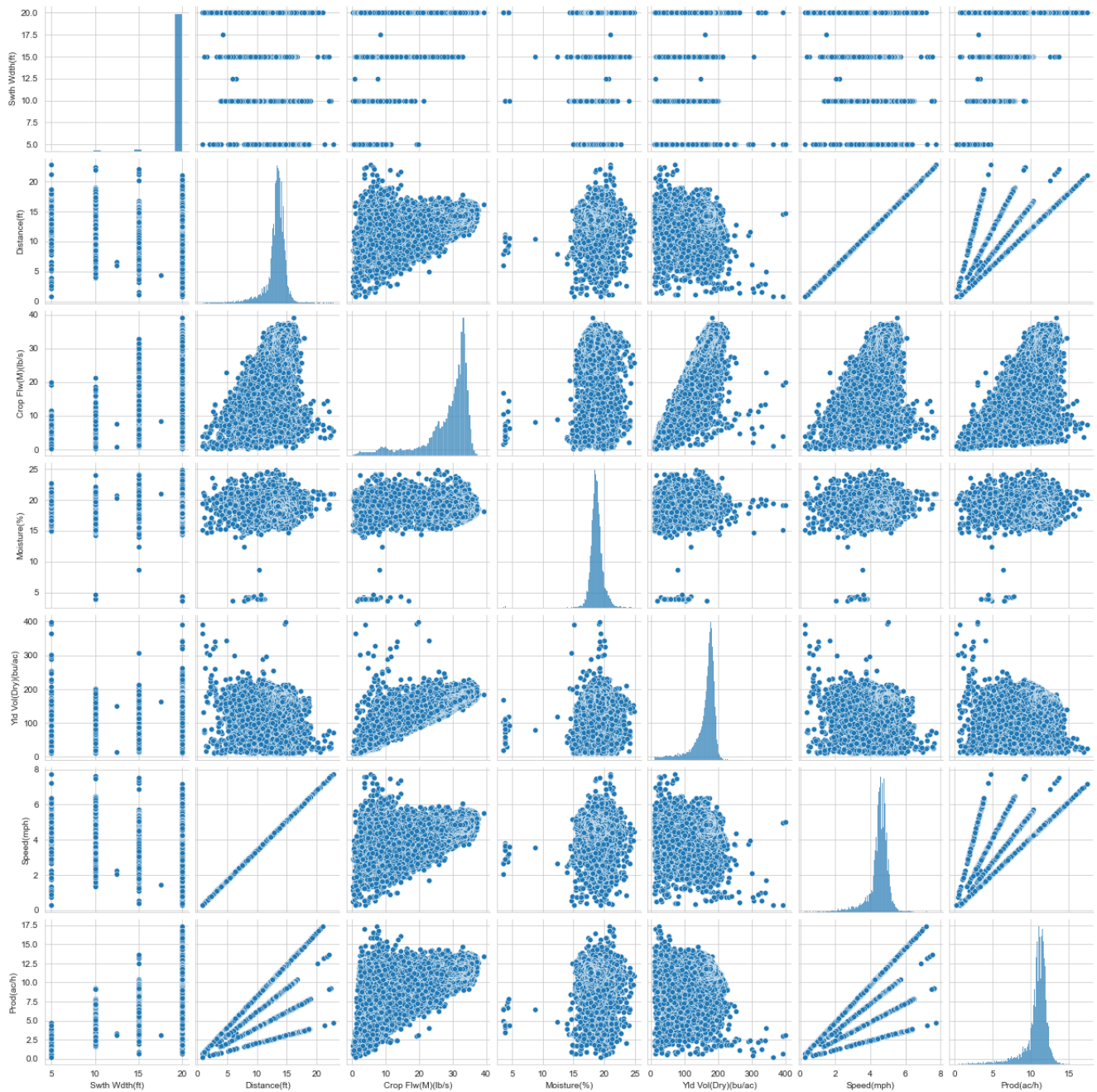


Figure 4. Pair plot with respect to all features.

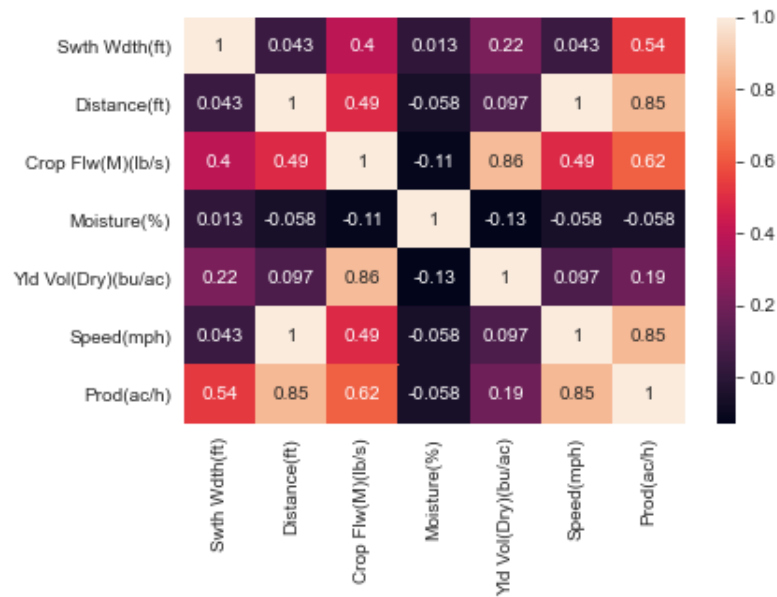


Figure 5. Correlation matrix heatmap for the data set.

Here we used the keras.callback “EarlyStopping” and monitored the validation loss. We use “EarlyStopping” as it terminates training when the chosen performance measure stops improving. As a result, if we look into the model accuracy plot (Figure 6) we observe that learning stops once the epoch number is 2. We also see that our model performs well enough with such a smaller epoch number and does not crudely over-fit or under-fit the training and testing data. Accuracy at that time is above 97%.

The line plot for the “model loss” (Figure 7) shows that the model is good at minimizing the loss function with fewer epochs.

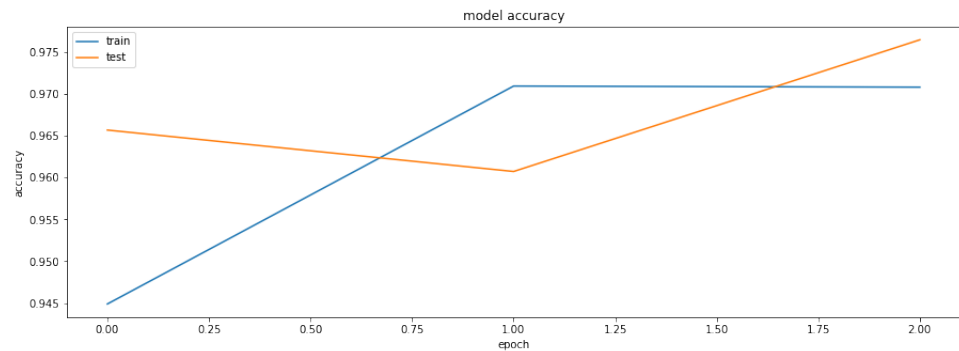


Figure 6. Model accuracy.

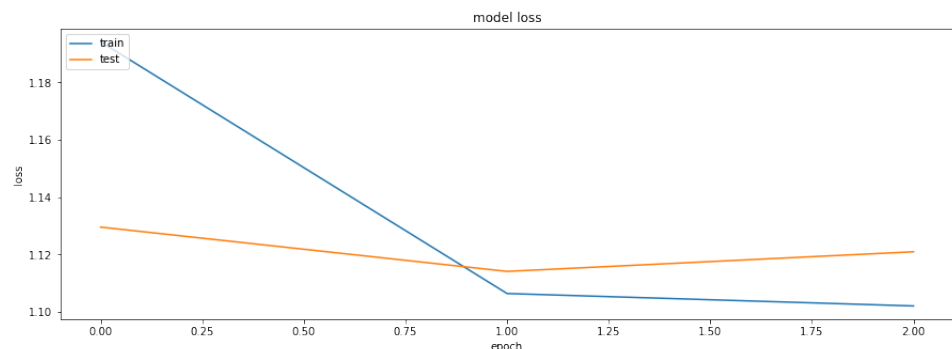


Figure 7. Model loss.

Method 2: The first part of this method is the same as in Method 1. However, instead of considering a , b , and c separately (as done in Method 1), for this method, we consider those simultaneously. In this case, with the same neural network as described in Method 1, we find that for the test data, the model correctly predicts 97.06% of the time. We create learning curves for different learning rates (lr) in Figures 8–13. When $lr = 1, 0.01$, we can observe from the loss curve and accuracy curve that the model performs poorly when fitting the training and testing dataset. The test data are over-fitted, and rather than showing any decay, the loss curve shows an increase as the epoch number increases. When $10^{-5} \leq lr < 10^{-6}$, the model performance is much more improved. This can be justified, as in this case, both model loss and model accuracy have a nicely fitted training and testing data set. When $lr > 10^{-6}$, the model performs much poorly, suggesting that there is probably no learning happening at all. The classification report for Method 1 and the model’s ability to correctly predict the test data set in Method 2 tells us that the model is not perfect and there is room for improvement. In our future work, we plan to improve our procedure and further polish this model to enhance the predictive ability for both these methods.

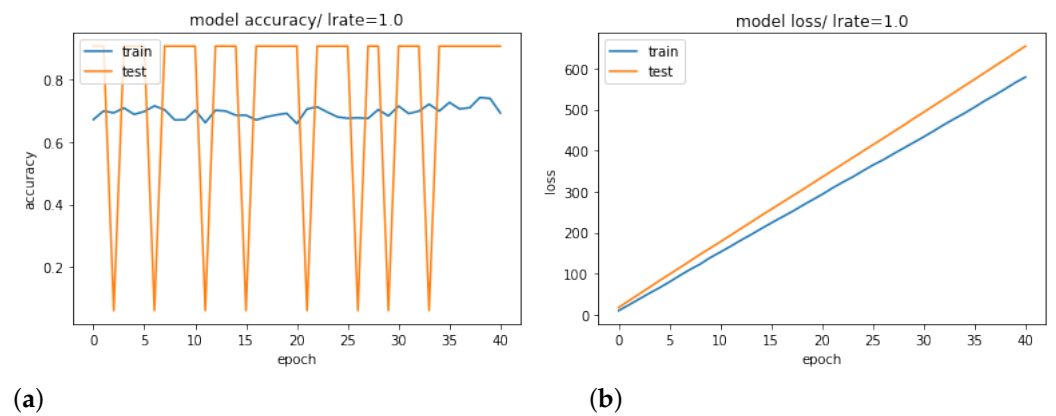


Figure 8. $lr = 1$. (a) Train accuracy: 0.911. (b) Test accuracy: 0.908.

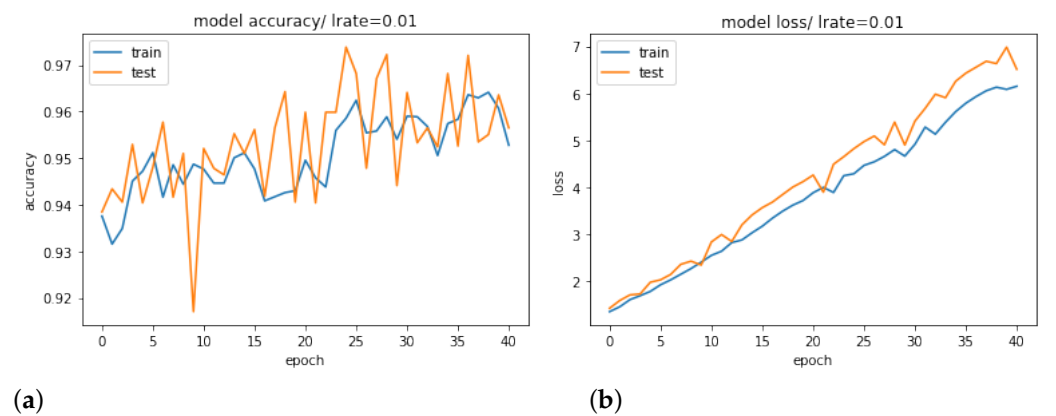


Figure 9. $lr = 10^{-2}$. (a) Train accuracy: 0.958. (b) Test accuracy: 0.956.

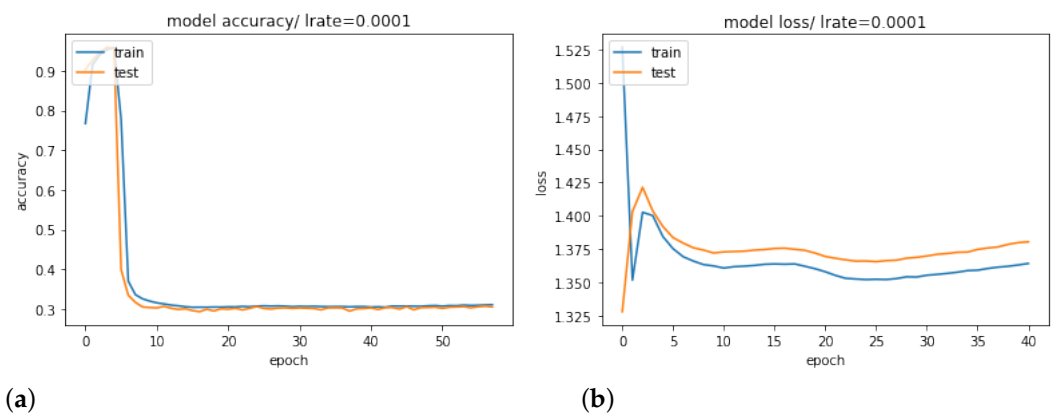


Figure 10. $lr = 10^{-4}$. (a) Train accuracy: 0.310. (b) Test accuracy: 0.305.

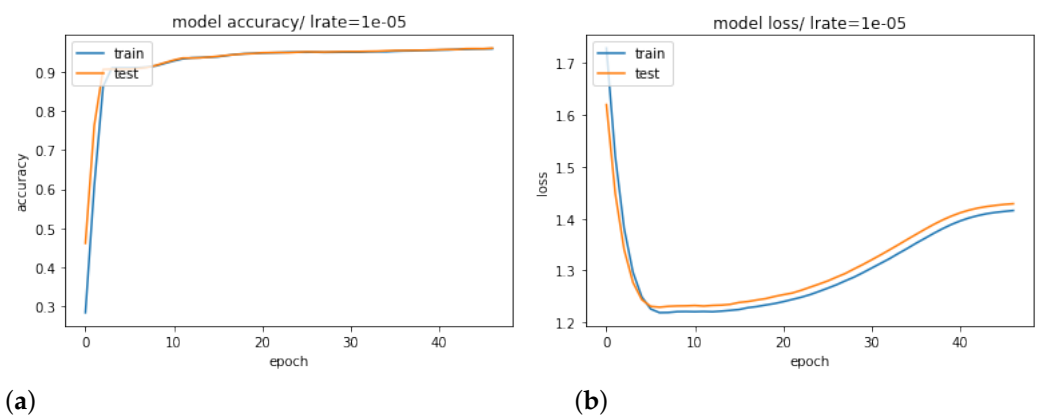


Figure 11. $lr = 1 \times 10^{-5}$. (a) Train accuracy: 0.961. (b) Test accuracy: 0.962.

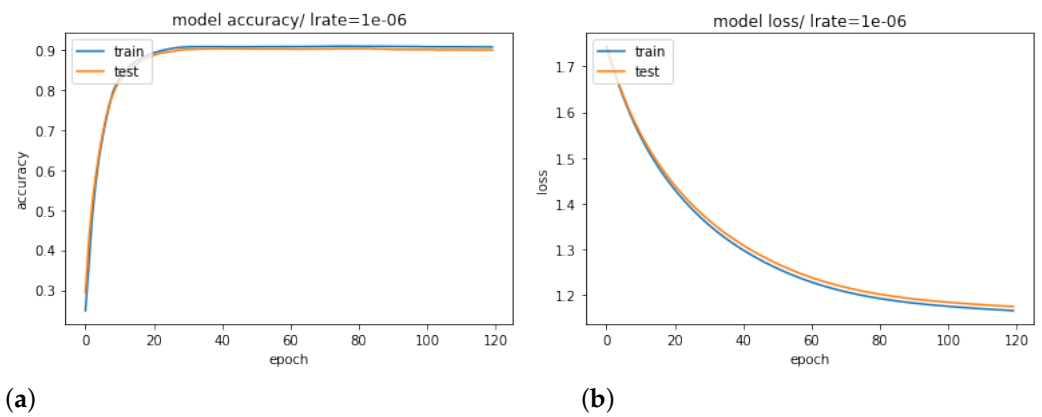


Figure 12. $lr = 1 \times 10^{-6}$. (a) Train accuracy: 0.907. (b) Test accuracy: 0.899.

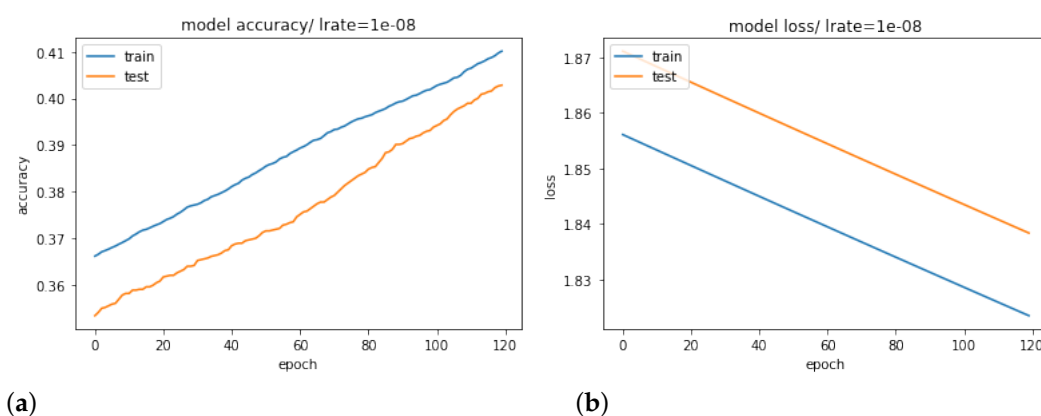


Figure 13. $lr = 1 \times 10^{-8}$. (a) Train accuracy: 0.410. (b) Test accuracy: 0.403.

From the analysis, it is clear that (4) can be an appropriate model for the *target variable*. It is also clear that for both the methods, the neural network model provides a very appropriate prediction of the *target variable*, with the help of *feature variables*.

4. Discussion

This paper analyzed yield data from a representative corn field in the upper midwestern United States. Yield data and data on other variables were collected every two seconds as the corn was harvested. These data are modeled to better understand corn yield in this field. In our data analysis, we wanted to answer the question: “Given intervals of fixed length, can a machine-learning-driven neural network model help us predict the interval on which sample points of feature Yld Vol (Dry)(bu/ac) belong to? And if so, how accurately does the model predict that correct interval?” In order to answer these questions, our first task was to create intervals (of fixed size) using the feature Yld Vol (Dry) (bu/ac) where the left end-points were kept as part of the interval. Once these intervals were constructed, we assigned the sample points of this feature Yld Vol (Dry) (bu/ac) into the correct interval. We labeled them with either 000 or 001 or 010 and so on (eight combination/labels) and created separate columns for these labels in our data set.

After this preparation stage was complete, we trained the data set with six features (Swth Wdth, Distance, Crop Flw, Moisture, Speed and Prod) where the label columns were treated as our test data. To avoid a highly complex deep neural network, we kept the number of hidden layers of our neural network to three. We compiled the model using the Adam optimizer and monitored loss using ‘categorical cross entropy’. When fitting our model, we used callback ‘EarlyStopping’ that monitored the validation loss. ‘EarlyStopping’ was implemented to speed up model training and to stop model fitting when the training data points is no longer learning efficiently; this also helps avoid overfitting. To observe that the learning does not become inefficient and to avoid overfitting, we computed validation loss and accuracy with different learning rates and analyzed them visually. In the last phase of our analysis, we obtained a classification report for our prediction. The final step of our analysis involved investigating how well our model correctly predicted the intervals. When compared to the test data points, the model correctly predicted almost 97% of the sample test points.

In [Stelzer and Barndorff-Nielsen \(2013\)](#), the authors introduce and analyze a multivariate supOU stochastic volatility (SV) model where they present an example of long memory in log returns in the SV supOU model, and in [Willinger et al. \(1999\)](#), the authors investigate whether stock price returns exhibit long-range dependence. Their work motivates us to investigate further with our pricing model to observe if our model exhibits long-range dependence as a stylized factor in our future project. At the same time, an in-depth analysis to determine whether our time series data exhibits any long-range dependence is something we wish to incorporate in a future sequel of our current work. Even though we have a model that could efficiently predict the interval on which sample data points of Yld Vol

(Dry) (bu/ac) belong to, there is still room for improvements and adjustments. Before training our data set, we can apply feature selection along with a random forest classifier to help us decide which features are more important than others. We can later replace them with the six features that we manually selected for our training purpose. Once we have allowed the machine to select the important features for us, in our modeling phase we can focus more on hyperparameter tuning by investigating with different epochs, batch sizes, and increasing the weights in the hidden layers.

Author Contributions: All the authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by USDA ARS grant number 58-6064-8-023.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from a cooperating farm manager and are available from the authors pending the permission of that individual.

Acknowledgments: The authors would like to thank the anonymous reviewers for their careful reading of the manuscript and for suggesting points to improve the quality of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Addey, Kwame Asiam, Saleem Shaik, William Nganje, and Indranil SenGupta. 2021. Implications of the Dirichlet processes mixture model on U.S. crop yield predictions in the presence of random shocks. submitted.
- Barndorff-Nielsen, Ole Eiler, and Niel Shephard. 2001a. Modelling by Lévy processes for financial econometrics. In *Lévy Processes: Theory and Applications*. Edited by Ole Eiler Barndorff-Nielsen, Thomas Mikosch and Sidney Resnick. Basel: Birkhäuser, pp. 283–318.
- Barndorff-Nielsen, Ole Eiler, and Neil Shephard. 2001b. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63: 167–241. [\[CrossRef\]](#)
- Bauer, Alan, Aaron George Bostrom, Joshua Ball, Christopher Applegate, Tao Cheng, Stephen Laycock, Sergio Moreno Rojas, Jacob Kirwan, and Ji Zhou. 2019. Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: A case study of lettuce production. *Horticulture Research* 6: 70. [\[CrossRef\]](#) [\[PubMed\]](#)
- Brunelli, Davide, Andrea Albanese, Donato d'Acunto, and Matteo Nardello. 2019. Energy Neutral Machine Learning Based IoT Device for Pest Detection in Precision Agriculture. *IEEE Internet of Things Magazine* 2: 10–13. [\[CrossRef\]](#)
- Chlingaryan, Anna, Salah Sukkarieh, and Brett Whelan. 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture* 151: 61–69. [\[CrossRef\]](#)
- Cont, Rama, and Peter Tankov. 2004. *Financial Modelling with Jump Processes*. CRC Financial Mathematics Series. Boca Raton: Chapman and Hall.
- Habtemicael, Semere, and Indranil SenGupta. 2016. Pricing variance and volatility swaps for Barndorff-Nielsen and Shephard process driven financial markets. *International Journal of Financial Engineering* 3: 1650027. [\[CrossRef\]](#)
- Hornng, Gwo-Jiun, Min-Xiang Liu, and Chao-Chun Chen. 2020. The Smart Image Recognition Mechanism for Crop Harvesting System in Intelligent Agriculture. *IEEE Sensors Journal* 20: 2766–81. [\[CrossRef\]](#)
- Ihsan, Atif, and Indranil SenGupta. 2018. Moments of the asset price for the Barndorff-Nielsen and Shephard model. *Lithuanian Mathematical Journal* 58: 408–20. [\[CrossRef\]](#)
- Issaka, Aziz, and Indranil SenGupta. 2017. Analysis of variance based instruments for Ornstein-Uhlenbeck type models: Swap and price index. *Annals of Finance* 13: 401–34. [\[CrossRef\]](#)
- Lemley, Joe, Shabab Bazrafkan, and Peter Corcoran. 2017. Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision. *IEEE Consumer Electronics Magazine* 6: 48–56. [\[CrossRef\]](#)
- Nicolato, Elisa, and Emmanouil Venardos. 2003. Option Pricing in Stochastic Volatility Models of the Ornstein-Uhlenbeck type. *Mathematical Finance* 13: 445–66. [\[CrossRef\]](#)
- Øksendal, Bernt, and Agnès Sulem-Bialobroda. 2007. *Applied Stochastic Control of Jump Diffusions*. Berlin/Heidelberg: Springer.
- Roberts, Michael, and Indranil SenGupta. 2020. Sequential hypothesis testing in machine learning, and crude oil price jump size detection. *Applied Mathematical Finance* 27: 374–95. [\[CrossRef\]](#)
- SenGupta, Indranil, William Wilson, and William Nganje. 2019. Barndorff-Nielsen and Shephard model: Oil hedging with variance swap and option. *Mathematics and Financial Economics* 13: 209–26. [\[CrossRef\]](#)

-
- Sharma, Abhinav , Arpit Jain, Prateek Gupta, and Vinay Chowdary. 2021. Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access* 9: 4843–73. [CrossRef]
- Shoshi, Humayra, and Indranil SenGupta. 2021. Hedging and Machine Learning Driven Crude Oil Data Analysis Using a Refined Barndorff–Nielsen and Shephard Model. *International Journal of Financial Engineering* 2021: 2150015. [CrossRef]
- Stelzer, Robert, and Ole Eiler Barndorff-Nielsen. 2013. The multivariate supOU stochastic volatility model. *Mathematical Finance* 23: 296.
- The International Society of Precision Agriculture. n.d. Available online: <https://www.ispag.org/> (accessed on 20 August 2021).
- Treboux, Jérôme, and Dominique Genoud. 2019. High Precision Agriculture: An Application Of Improved Machine-Learning Algorithms. Paper presented at 2019 6th Swiss Conference on Data Science (SDS), Bern, Switzerland, June 14; pp. 103–8. [CrossRef]
- Willinger, Walter, Murad S. Taqqu, and Vadim Teverovsky. 1999. Stock market prices and long range dependence. *Finance and Stochastics* 3: 1–13. [CrossRef]
- Wilson, William, William Nganje, Semere Gebresilasie, and Indranil SenGupta. 2019. Barndorff–Nielsen and Shephard model for hedging energy with quantity risk. *High Frequency* 2: 202–14. [CrossRef]