EVALUATION OF PRICING WIND SPEED DERIVATIVE CONTRACTS

DIANA SON

ABSTRACT. This paper acts as the final project for the Quantitative Capstone in Environmental Finance course at Florida International University, under the supervision of Dr. Pablo Olivares and Dr. Enrique Villamor, for Spring 2026.

CONTENTS

1

1. INTRODUCTION

Background
Motivation

Definition 1 (Standard Brownian Motion)**.** A standard Brownian motion $\{v_t\}_{t\geq 0}$ is a continuous in time stochastic process satisfying the following conditions:

- $v_0 = 0$ almost surely,
- v has **stationary increments**, that is, for s, t with $s, t \in [0, \infty)$, the distribution of $v_t - v_s$ is the same as the distribution of v_{t-s},
- v has **independent increments**, that is, for $t_1, t_2, ..., t_n \in [0, \infty)$ with $t_1 < t_2 < ... < t_n$, the random variables $v_{t_1}, v_{t_2} - v_{t_1}, ..., v_{t_n} - v_{t_{n-1}}$ are independent,
- v_t is normally distributed with mean 0 and variance t for each $t \in (0, \infty)$.
- with probability 1, $t \mapsto v_t$ is continuous on $[0, \infty)$,

and with the standard multiplication rules:

$$dt \cdot dt = 0, \quad dv_t \cdot dt = 0, \quad dv_t \cdot dv_t = dt.$$

Date: January 2026.

[1]**Keywords:** stochastic calculus, PDE theory, numerical analysis, wind speed derivative pricing

Definition 2 (Ornstein-Ulhenbeck Process with Seasonal Mean)**.** A process $\{X_t\}_{t\geq 0}$ follows an Ornstein-Ulhenbeck process with seasonal mean if it satisfied the SDE:

$$dX_t = \kappa(\mu(t) - X_t)dt + \sigma dW_t \tag{1}$$

where the long-run mean $\mu(t)$ is a deterministic seasonal function of time, specifically as a truncated Fourier series:

$$\mu(t) = \theta_1 + \theta_2 \sin(\omega t) + \theta_3 \cos(\omega t) \tag{2}$$

where

- θ_1 is the overall mean level
- θ_2 and θ_3 controls the amplitude and phase of the seasonal cycle
- $\omega = \frac{2\pi}{365.25}$ is the angular frequency for an annual cycle.

Then, the key properties of the SDE are:

- $\kappa > 0$ is the mean reversion speed, which ensures that X_t continuously reverts towards the time varying seasonal mean $\mu(t)$,
- $\sigma > 0$ is the volatility of the process,
- W_t is the standard Brownian motion,
- and the conditional mean and variance are given by:

$$E\left[X_t|X_s\right] - X_s e^{-\kappa(t-s)} + \int_s^t \kappa\mu(u)e^{-\kappa(t-u)}du$$

$$\mathrm{Var}(X_t|X_s) = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa(t-s)})$$

Lemma 1.1 (Itô's Lemma)**.** *Let X_t satisfy the SDE:*

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t$$

and let $f(X_t, t)$ be a twice continuous, differentiable function. Then by itô's Lemma:

$$df = \left(\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial X} + \frac{1}{2}b^2\frac{\partial^2 f}{\partial X^2}\right)dt + b\frac{\partial f}{\partial X}dW_t \tag{3}$$

Definition 3 (No-Arbitrage Condition)**.** The no-arbitrage condition states that the discounted price process of any traded asset must be a martingale under a risk neutral probability measure $\mathbb{Q}$, then, for a derivative with price V_t, the discounted process $e^{-rt}V_t$ must satisfy

$$e^{-rt}V_t = E^{\mathbb{Q}}\left[e^{-rT}V_T|\mathcal{F}_t\right] \quad \equiv \quad V_t = e^{-r(T-t)}E^{\mathbb{Q}}\left[V_T|\mathcal{F}_t\right]$$

That is, today's fair price of a derivative is the expected future payoff, discounted back to present-day at the risk-free rate r.

Theorem 1.2 (Feynman-Kac Theorem)**.** *Let X_t satisfy the SDE:*

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t$$

and suppose $D(X_t, t)$ satisfies the terminal value problem:

$$\frac{\partial D}{\partial t} + a(X_t, t)\frac{\partial D}{\partial x} + \frac{1}{2}b(X_t, t)\frac{\partial^2 D}{\partial x^2} - rD = 0$$

with terminal condition

$$D(X_T, T) = \Phi(X_T)$$

where $\Phi(X_T)$ is the payoff function of the derivative at maturity time T. Then by the Feynman-Kac Theorem, the solution to this PDE with terminal condition has a probabilistic representation:

$$D(X_t, t) = e^{-r(T-t)}E^{\mathbb{Q}}\left[\Phi(X_T)|\mathcal{F}_t\right].$$

2. Derive PDE from SDE

The dynamics of wind speed is modeled by the following stochastic differential equation:

$$\begin{aligned} dX_t &= (\mu_X(t) - \lambda_X X_t)dt + \sigma_X dv_t \\ \mu_X(t) &= \theta_1 + \theta_2 \sin(\omega t) + \theta_3 \cos(\omega t) \end{aligned} \tag{4}$$

where X_t, the daily average wind speed (in m/s), follows a mean-reverting process in continuous time with seasonality μ_X, $\lambda_X > 0$ is the mean reversion speed, ω_X is the volatility of wind speed, v_t is the standard Brownian motion, and the wind speed derivative $V(X_t, t)$ is written on this process.

The model in equation (4) is an Ornstein-Ulhenbeck process (1) with seasonal mean, capturing the annual cycle of wind speed.

To derive the associate partial differential equation from the stochastic differential equation (4), we first apply Itô's Lemma to $V(X_t, t)$ and simplify using Itô's multiplication table:

$$dV = \left[\frac{\partial V}{\partial t} + (\mu_X(t) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2}\right] dt + \sigma_X \frac{\partial V}{\partial x} dv_t.$$

That is, over a small time interval, V changes by a predictable, deterministic amount plus a random shock.

Now, consider the discounted derivative price $U_t = e^{-rt}V(X_t, t)$. Now, apply Itô's Lemma to U_t:

$$dU_t = e^{-rt}\left[\left(-rV + \frac{\partial V}{\partial t} + (\mu_X(t) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 V}{\partial x^2}\right) dt + \sigma \frac{\partial V}{\partial x} dv_t\right].$$

the no-arbitrage condition requires that $U_t = e^{-rt}V_t$ is a martingale under $\mathbb{Q}$, meaning that the dt term must equal 0. Thus, we have the pricing PDE:

$$\frac{\partial V}{\partial t} + (\mu_X(t) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2} - rV = 0 \tag{5}$$

with the contract payoff as the terminal condition. Hence, equation (5) is backwards in time $t : T \to 0$.

For numerical convenience, it's beneficial to convert this terminal value problem into an initial value problem via a simple change of variables.

Let $\tau = T - t$, then the derivatives transform as:

$$\frac{\partial}{\partial t} = -\frac{\partial}{\partial \tau}.$$

Note, the spatial derivatives do not change. Then, substitute back into equation (5):

$$-\frac{\partial V}{\partial \tau} + (\mu_X(T-\tau) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2} - rV = 0.$$

Thus, we've obtained the initial value problem:

$$\begin{aligned} &\frac{\partial V}{\partial \tau} = (\mu_X(T-\tau) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2} - rV \\ &\mu_X(T-\tau) = \theta_1 + \theta_2 \sin(\omega(T-\tau)) + \theta_3 \cos(\omega(T-\tau)) \\ &V(X, 0) = \Phi(X) \end{aligned} \tag{6}$$

where $\tau \in [0, T]$ is the time since maturity.

Remark 2.1 (Different Initial Conditions)**.** The initial condition to the initial value problem depends on the contract being priced. In the context of wind speed derivatives, the common initial conditions include:

- call option: $\Phi(X) = \max(X_T - K, 0)$, which is the initial condition I will be using in the upcoming sections, pays off when the wind speed exceeds the strike wind speed K

- put option: $\phi(X) = \max(K - X_T, 0)$ pays off when the wind speed falls below the strike wind speed K.
- forward contract: $\Phi(X) = X_T - K$ is the linear payoff on wind speed at maturity.
- binary option: $\Phi(X) = 1_{\{X_T \geq K\}}$ pays one unit if wind speed exceeds K and 0 otherwise.
- capped payoff: $\Phi(X) = \min(\max(X_T - K, 0), C)$ is a like a call option but the maximum payoff is capped at C.

Remark 2.2 (Different Drifts)**.** For the purposes of this project, the model uses the physical drift $\mu_X(t) - \lambda_X x$, however, other possible drifts to consider include:

- Risk-neutral drift:
$$\mu_X(t) - \lambda_X x - \sigma_X \rho_X \phi$$
where $\sigma_X \rho_x \phi$ is the risk adjustment that's coming from the wind speed risk that could be partially hedge by the stock market.
- Drift with constant risk premium:
$$\mu_X(t) - \lambda_X x - \gamma \sigma_x$$
where σ_X is constant and γ is the constant market price of wind speed. Note the sign of γ indicates whether investors demand compensation or pay a premium.
- Good-deal lower/upper bounds drift:
$$\mu_X(t) - \lambda_X x \pm \sqrt{A^2 - \phi^2}\sigma_X \sqrt{1 - \rho_X^2}$$
gives the lower $(-)$, interpret as buyer's price, and upper $(+)$, interpret as seller's price, boundaries, producing a price interval that would be consistent with the no-arbitrage in the incomplete market.

3. Numerical Schemes

In this section, the numerical schemes used to solve the initial value problem to the pricing PDE is detailed step-by-step. We solve the associate PDE with 3 different schemes to later compare the accuracy, stability, and convergence of each method.

Start with the complete initial value problem:

$$\begin{cases} \frac{\partial V}{\partial \tau} = (\mu_X(T-\tau) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2} - rV, \quad x \in [a,b] \text{ and } \tau \in [0,T] \\ \mu_X(T-\tau) = \theta_1 + \theta_2 \sin(\omega(T-\tau)) + \theta_3 \cos(\omega(T-\tau)) \\ V(X,0) = \Phi(X) = \max(X-K, 0) \\ V(a,\tau) = 0, \;\; \forall \tau \in [0,T] \\ \frac{\partial^2 V}{\partial x^2}(b,\tau) = 0, \;\; \forall \tau \in [0,T] \end{cases} \tag{7}$$

with left Dirchlet boundary condition and right linear boundary condition are the standard choice for this type of problem. To numerically solve the PDE (7), first discretize the domains into a uniform grid and use the common notation for time and spatial discretizations:

$$x_j = a + j\Delta x, \quad j = 0, ..., M,$$

$$\tau_n = n\Delta\tau, \qquad n = 0, ..., N.$$

Then, we approximate the first and second derivatives with the following standard finite difference:

Forward Difference:
$$\frac{\partial f}{\partial t} \approx \frac{f_j^{n+1} - f_j^n}{\Delta t}$$

Central Difference:
$$\frac{\partial f}{\partial x} \approx \frac{f_{j+1}^n - f_{j-1}^n}{\Delta x}$$

2nd-order Central Difference:
$$\frac{\partial^2 f}{\partial x^2} \approx \frac{f_{j+1}^n - 2f_j^n + f_{j-1}^n}{(\Delta x)^2}$$

where $f_j^n = f(\tau_n, x_j)$. The first scheme we considered is the explicit *forward* Euler scheme.

3.1. **Explicit Euler.** To derive the explicit Euler scheme for the pricing PDE (7), substitute into (7) the finite difference approximations at time n, then rearrange the equation explicitly for V_j^{n+1}:

$$\begin{aligned}
\frac{\partial V}{\partial \tau} &= (\mu_X(T-\tau) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2} - rV \\
\frac{V_j^{n+1} - V_j^n}{\Delta\tau} &= -rV_j^n + a_j^n \frac{V_{j+1}^n - V_{j-1}^n}{2\Delta x} + \frac{1}{2}\sigma_X^2 \frac{V_{j+1}^n - 2V_j^n + V_{j-1}^n}{(\Delta x)^2} \\
V_j^{n+1} &= V_j^n + \Delta\tau\left[-rV_j^n + a_j^n \frac{V_{j+1}^n - V_{j-1}^n}{2\Delta x} + \frac{1}{2}\sigma_X^2 \frac{V_{j+1}^n - 2V_j^n + V_{j-1}^n}{(\Delta x)^2}\right]
\end{aligned}$$

where $a_j^n = \mu_X(T-\tau_n) - \lambda_X x_j$. Note that the explicit Euler scheme is stable if and only if the stability condition

$$\frac{\sigma_X^2 \Delta\tau}{(\Delta x)^2} \leq \frac{1}{2}$$

holds. This condition can be derived explicitly via the Von Neumann Stability Analysis procedure.

The second scheme we consider is the implicit *backward* Euler scheme.

3.2. **Implicit Euler.** To derive the implicit Euler scheme for the initial value problem (7), substitute into the parabolic PDE the finite difference approximations at time $n+1$ and rearrange:

$$\begin{aligned}
\frac{\partial V}{\partial \tau} &= (\mu_X(T-\tau) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2} - rV \\
\frac{V_j^{n+1} - V_j^n}{\Delta\tau} &= -rV_j^{n+1} + a_j^{n+1} \frac{V_{j+1}^{n+1} - V_{j-1}^{n+1}}{2\Delta x} + \frac{1}{2}\sigma_X^2 \frac{V_{j+1}^{n+1} - 2V_j^{n+1} + V_{j-1}^{n+1}}{(\Delta x)^2}
\end{aligned}$$

where $a_j^{n+1} = \mu_X(T-\tau_{n+1}) - \lambda_X x_j$.
Now, rearrange equation into the form $AV_j^{n+1} = V_j^n$ where $A^{(M+1)\times(M+1)}$ is a tridiagonal matrix:

$$\underbrace{\left[-\Delta\tau\left(\frac{\sigma_X^2}{2(\Delta x)^2} - \frac{a_j^{n+1}}{2\Delta x}\right)\right]}_{\text{lower diagonal } l_j} V_{j-1}^{n+1} + \underbrace{\left[1 + \Delta\tau\left(r + \frac{\sigma_X^2}{(\Delta x)^2}\right)\right]}_{\text{diagonal } d_j} V_j^{n+1} + \underbrace{\left[-\Delta\tau\left(\frac{a_j^{n+1}}{2\Delta x} + \frac{\sigma_X^2}{2(\Delta x)^2}\right)\right]}_{\text{upper diagonal } u_j} V_{j+1}^{n+1} = V_j^n$$

Next, simplify the equation further by rewriting into matrix representation. Note, this allows for clearer numerical implementation of the implicit Euler scheme. Thus, we have:

$$\underbrace{\begin{pmatrix}
1 & 0 & 0 & 0 & \cdots & 0 \\
l_1 & d_1 & u_1 & 0 & \cdots & 0 \\
0 & l_2 & d_2 & u_2 & \cdots & 0 \\
\vdots & & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & l_{M-1} & d_{M-1} & u_{M-1} \\
0 & \cdots & 0 & 1 & -2 & 1
\end{pmatrix}}_{\text{tridiagonal matrix A}}
\begin{pmatrix} V_0^{n+1} \\ V_1^{n+1} \\ V_2^{n+1} \\ \vdots \\ V_{M-1}^{n+1} \\ V_M^{n+1} \end{pmatrix}
=
\begin{pmatrix} 0 \\ V_1^n \\ V_2^n \\ \vdots \\ V_{M-1}^n \\ 0 \end{pmatrix}
\implies AV_j^{n+1} = V_j^n \implies V_j^{n+1} = A \setminus V_j^n$$

Note, implicit Euler is unconditionally stable and the proof can be shown via Von Neumann Stability Analysis. The third and last scheme we consider is the Crank-Nicolson scheme.

3.3. **Crank-Nicolson.** Next, to derive the crank-Nicolson scheme, substitute the finite difference approximations into the PDE by average the RHS between times n and $n+1$, effectively implementing both explicit and implicit Euler schemes. Then, we have:

$$\frac{\partial V}{\partial \tau} = (\mu_X(T-\tau) - \lambda_X x)\frac{\partial V}{\partial x} + \frac{1}{2}\sigma_X^2 \frac{\partial^2 V}{\partial x^2} - rV$$

$$\frac{V_j^{n+1} - V_j^n}{\Delta\tau} = \frac{1}{2}\underbrace{\left[-rV_j^n + a_j^n \frac{V_{j+1}^n - V_{j-1}^n}{2\Delta x} + \frac{1}{2}\sigma_X^2 \frac{V_{j+1}^n - 2V_j^n + V_{j-1}^n}{(\Delta x)^2}\right]}_{\text{explicit Euler}} + \frac{1}{2}\underbrace{\left[-rV_j^{n+1} + a_j^{n+1} \frac{V_{j+1}^{n+1} - V_{j-1}^{n+1}}{2\Delta x} + \frac{1}{2}\sigma_X^2 \frac{V_{j+1}^{n+1} - 2V_j^{n+1} + V_{j-1}^{n+1}}{(\Delta x)^2}\right]}_{\text{implicit Euler}}$$

Simplify the expression by rearranging terms to achieve n and $n+1$ terms on opposite sides:

$$\underbrace{\left[\frac{\Delta\tau}{2}\left(\frac{a_j^{n+1}}{2\Delta x} - \frac{\sigma_X^2}{2(\Delta x)^2}\right)\right]}_{\text{lower diagonal of A}} V_{j-1}^{n+1} + \underbrace{\left[1 + \frac{\Delta\tau}{2}\left(r + \frac{\sigma_X^2}{2(\Delta x)^2}\right)\right]}_{\text{diagonal of A}} V_j^{n+1} + \underbrace{\left[-\frac{\Delta\tau}{2}\left(\frac{a_j^{n+1}}{2\Delta x} + \frac{\sigma_X^2}{2(\Delta x)^2}\right)\right]}_{\text{upper diagonal of A}} V_{j+1}^{n+1}$$

$$= \underbrace{\left[\frac{-\Delta\tau}{2}\left(\frac{a_j^{n}}{2\Delta x} - \frac{\sigma_X^2}{2(\Delta x)^2}\right)\right]}_{\text{lower diagonal of B}} V_{j-1}^{n} + \underbrace{\left[1 - \frac{\Delta\tau}{2}\left(r + \frac{\sigma_X^2}{2(\Delta x)^2}\right)\right]}_{\text{diagonal of B}} V_j^{n} + \underbrace{\left[\frac{\Delta\tau}{2}\left(\frac{a_j^{n}}{2\Delta x} + \frac{\sigma_X^2}{2(\Delta x)^2}\right)\right]}_{\text{upper diagonal of B}} V_{j+1}^{n}$$

where $A^{(M+1)\times(M+1)}$ and $B^{(M+1)\times(M+1)}$ are tri-diagonal matrices. Next, simplify by rewriting into matrix representation:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ l_1^A & d_1^A & u_1^A & 0 & \cdots & 0 \\ 0 & l_2^A & d_2^A & u_2^A & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & l_{M-1}^A & d_{M-1}^A & u_{M-1}^A \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}}_{\text{tri-diagonal matrix A}} \begin{pmatrix} V_0^{n+1} \\ V_1^{n+1} \\ V_2^{n+1} \\ \vdots \\ V_{M-1}^{n+1} \\ V_M^{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ l_1^B & d_1^B & u_1^B & 0 & \cdots & 0 \\ 0 & l_2^B & d_2^B & u_2^B & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & l_{M-1}^B & d_{M-1}^B & u_{M-1}^B \\ 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix}}_{\text{tri diagonal matrix B}} \begin{pmatrix} 0 \\ V_1^{n} \\ V_2^{n} \\ \vdots \\ V_{M-1}^{n} \\ 0 \end{pmatrix}$$

Thus, the Crank-Nicolson scheme is simplified down to:

$$AV_j^{n+1} = BV_j^n \implies V_j^{n+1} = A \setminus (BV_j^n)$$

Note, like implicit Euler, the Crank-Nicolson scheme is unconditionally stable and can be proven as such via Von Neumann Stability Analysis calculations.

4. Evaluating Initial Parameters

In this section, the goal is the understand and estimate the parameters that impact the initial value problem (7). In the model (4), the daily average wind speed X_t in (m/s) is modeled by a Ornstein-Ulhenbeck process with deterministic seasonal mean

$$\mu_X(t) = \theta_1 + \theta_2 \sin(\omega t) + \theta_3 \cos(\omega t)$$

is a truncated Fourier series function with $\omega = \frac{2\pi}{365.25}$. The other parameters are

- $\lambda_X > 0$ is the mean reversion speed
- $\sigma_x > 0$ is the volatility, which measures the magnitude of random daily fluctuations
- θ_1 is the overall mean level of wind speed
- θ_2 and θ_3 are the seasonal coefficients controlling the amplitude and phase of the annual cycle.

Then, the focus of this section is to utilize Miami wind speed datasets to estimate the parameters listed above. Later, the estimated parameters will be used when numerically solving for V in (7).

Remark 4.1. The two datasets we're working with come from the NOAA Global Historical Climatology Network (GHCN), recorded at the Miami International Airport (MIA) station. Dataset 1 has daily observations from January 1st, 1990 to August 25th, 2024, while dataset 2 has daily observations from January 1st, 1981 to August 27th, 2023.

First, the seasonal mean function $\mu_X(t)$ is estimated by ordinary least squares regression.

4.1. **Ordinary Least Squares.** In mathematical statistics, the ordinary least squares (OLS) method is widely known and used. As an algebraic procedure, OLS fits linear equations to observable dataset by minimizing the sum of the squared errors between observed and fitted values. The original author of the OLS method is debated, some understand that OLS was developed by Legendre in 1805 [include citation] but Gauss claims to have developed OLS as early as 1795.

To understand ordinary least squares method, consider a linear regression model

$$y = A\theta + \epsilon$$

where $y \in \mathbb{R}^N$ is the vector of observations, $A \in \mathbb{R}^{N\times p}$ is the design matrix of known parameters, $\theta \in \mathbb{R}^p$ is the vector of unknown parameters, and $\epsilon \in \mathbb{R}^n$ is the vector of errors. Then the OLS estimator $\hat{\theta}^{\text{OLS}}$ minimizes the sum of squared residuals:

$$\hat{\theta}^{\text{ OLS}} = \arg\min_{\theta} \|y - A\theta\|^2 = \arg\min_{\theta} \sum_{t=1}^{N} (y_t - a_t^\top \theta)^2 \tag{8}$$

where $a_t^\top$ is the t-th row of A. Then, take the derivative of the objective w.r.t. θ and setting the result equal to 0 gives the normal equations:

$$A^\top A\hat{\theta} = A^\top y.$$

Provided that $A^\top A$ is invertible, then the unique closed-form solution is

$$\hat{\theta}^{\text{ OLS}} = (A^\top A)^{-1}A^\top y. \tag{9}$$

Then, $\mu_X(t)$ is a linear equation with 3 unknown coefficients $(\theta_1, \theta_2, \theta_3)$, given N daily observations $\{X_t\}_{t=1}^N$ with corresponding day-of-year indices $\{d_t\}_{t=1}^N$ (where $d_t \in \{1, ..., 365\}$), we can construct the design matrix A as:

$$A = \begin{pmatrix} 1 & \sin(\omega d_1) & \cos(\omega d_1) \\ 1 & \sin(\omega d_2) & \cos(\omega d_2) \\ \vdots & \vdots & \vdots \\ 1 & \sin(\omega d_N) & \cos(\omega d_N) \end{pmatrix} \in \mathbb{R}^{N\times 3}, \qquad \text{where } \omega = \frac{2\pi}{365.25} \tag{10}$$

Then, the OLS estimator is $\hat{\theta} = (A^\top A)^{-1}A^\top X$ where $X = (X_1, X_2, ..., X_N)^\top$, and minimizes

$$\hat{\theta} = \arg\min_{\theta_1,\theta_2,\theta_3} \sum_{t=1}^{N} (X_t - \theta_1 - \theta_2 \sin(\omega d_t) - \theta_3 \cos(\omega d_t))^2, \tag{11}$$

the sum of squared derivations of wind speed from the seasonal mean.

Remark 4.2. The seasonal function $\mu_X(t)$ is setup to be evaluated at the *day of the year* instead of the cumulative day index from the start of the dataset to produce a function that repeats annually, capturing the seasonal pattern.

The next step is to deseasonalize.

4.2. **Deseasonalize.** The wind speed series X_t is a combination of deterministic seasonal pattern (*smooth, predictable annual cycle*) and stochastic component (*the random day-by-day fluctuations, aka atmospheric variability*). The deseasonalize step will isolate the stochastic dynamics $\tilde{X}_t$ so that the next steps in the procedure to estimate the random mean-reverting behavior of $\tilde{X}_t$.

Once the seasonal mean parameters $(\theta_1, \theta_2, \theta_3)$ are estimated in the first step, compute the deseasonalized residual series by subtracting the estimated seasonal mean from the observed wind speed to isolate the stochastic part:

$$\tilde{X}_t = X_t - \hat{\mu}_X(t) = X_t - \theta_1 - \sin(\omega d_t) - \theta_3 \cos(\omega d_t), \qquad t = 1, ..., N.$$

Remark 4.3. After deseasonalizing, $\tilde{X}_t$ satisfies the simplified SDE

$$d\tilde{X}_t = -\lambda_X \tilde{X}_t dt + \sigma_X dv_t$$

with exact analytical solution, with $\Delta t = 1$ is:

$$\tilde{X}_t = e^{-\lambda_X \Delta t} \tilde{X}_{t-1} + \epsilon_t$$

where ϵ_t has zero mean and variance $\sigma_\epsilon^2 = \frac{\sigma_X^2}{2\lambda_X}(1 - e^{-2\lambda_X \Delta t})$.

Then, the third step is to compute the regression on the residuals.

4.3. **AR(1) Regression on the Residuals.** An AR(1), autoregressive model of order 1, is the discrete-time model where the current value has a linear dependence on the preceding value plus an independent shock:

$$\tilde{X}_t = \phi \tilde{X}_{t-1} + \epsilon_t, \qquad \epsilon_t \overset{\text{i.i.d.}}{\sim} N(0, \sigma_\epsilon^2), \tag{12}$$

where $\phi \in (-1, 1)$ is the autoregressive coefficient that measures how much of today's value depends on yesterday's value, ϵ_t is the independent Gaussian random shock. By regressing $\tilde{X}_t$ on $\tilde{X}_{t-1}$, we directly estimate ϕ, which is needed to analytical compute back λ_X and σ_X

Comparing the coefficient from (12) to the analytical solution from Remark 4.3, it follows that $\phi = e^{-\lambda_X \Delta t}$, that is, the AR(1) coefficient ϕ depends on the mean reversion speed λ_X of the continuous-time Ornstein-Ulhenbeck process. Note, since $\lambda_X > 0$, it follows that $\phi \in (0, 1)$.

Now, define vectors $Y = (\tilde{X}_2, \tilde{X}_3, ..., \tilde{X}_N)^\top$ and $Z = (\tilde{X}_1, \tilde{X}_2, ..., \tilde{X}_{N-1})^\top$ that represent the current and lagged values, respectively. Then, the AR(1) model in (12) is a linear regression of Y on Z with no intercept. Then, apply the OLS formula (9) with design matrix Z, the estimator simplifies to:

$$\hat{\phi} = \frac{Z^\top Y}{Z^\top Z} = \frac{\sum\limits_{t=2}^{N} \tilde{X}_{t-1} \tilde{X}_t}{\sum\limits_{t=2}^{N} \tilde{X}_{t-1}^2}, \tag{13}$$

which is the ratio of the lag-1 sample autocovariance to the sample variance of the deseasonalized residuals. For example, $\hat{\phi} = 0.6575$ means that on average 65.75% of today's wind speed deviation from the seasonal mean will continue to the next day.

Then, the AR(1) residuals are $\epsilon_t = \tilde{X}_t - \phi \tilde{X}_{t-1}$, $t = 2, .., N$ that represent the portion of today's wind speed deviation that are not explained by yesterday's value, and their variance is $\sigma_\epsilon^2 = \frac{1}{N-1} \sum\limits_{t=2}^{N} \hat{\epsilon}_t^2$.

Finally, the last step is to simply solve for the continuous-time parameters, λ_X and σ_X. From the AR(1) coefficient $\phi = e^{-\lambda_X \Delta t}$, λ_X can be solved for directly by taking the natural logarithm of both sides and get $\hat{\lambda}_X = -\ln \hat{\phi}$ where $\Delta t = 1$. Then, by the variance in Remark 4.3, σ_X can be solved to get

$$\hat{\sigma}_X = \sqrt{\frac{2\hat{\lambda}_X \hat{\sigma}_\epsilon^2}{1 - e^{-2\hat{\lambda}_X \Delta t}}}.$$

Thus, using the procedure detailed in this section, we can estimate the values of each parameters based on the 2 wind speed datasets recorded from Miami International Airport.

4.4. **Dataset 1.** Following the procedure detailed in this section, the parameters $\lambda_X, \sigma_x, \theta_1, \theta_2,$ and θ_3 are estimated based off the first dataset recorded at Miami International Airport with observations from January 1st, 1990 to August 25th, 2024, containing $12,654$ total observations with the following mean and standard deviation:

$$\text{mean speed} = 3.5361 \text{ (m/s)}, \qquad \text{std speed } = 1.4092 \text{ (m/s)}.$$

To better examine the estimated parameter values, Figure 1 shows the bar graphs of each estimated parameter value per year, where the solid red line is the parameter value calculated over the entire dataset and the dashed black line is mean of the per-year parameter values. For exact values, refer to Section 7 for the full table.

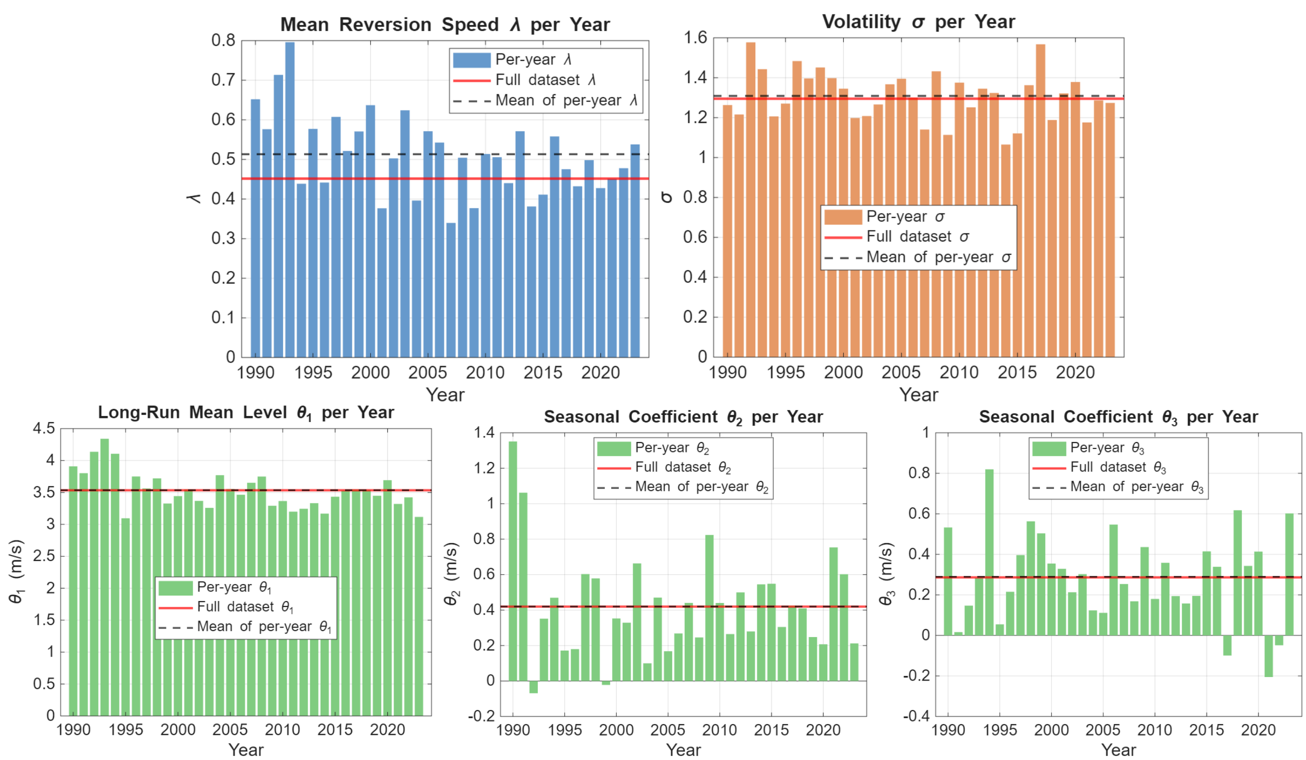

FIGURE 1. The bar plots of λ (top left), σ (top right), θ_1 (bottom left), θ_2 (bottom center), and θ_3 (bottom right) calculated values per year and over the entire dataset 1.

4.5. **Dataset 2.** Now, we apply the same procedure detailed in this section to dataset 2 containing wind speed observations record from Miami International Airport from January 1, 1984 to August 27th, 2023, totaling $14,482$ observations with the following mean and standard deviation:

$$\text{mean speed} = 3.7133 \text{ (m/s)}, \qquad \text{std speed } = 1.4220 \text{ (m/s)}.$$

Then, Figure 2 shows the corresponding bar graph plots to each parameter to better illustrate the variability of each parameter over the 40-year period. For exact values, refer to Section 7 for the full table.

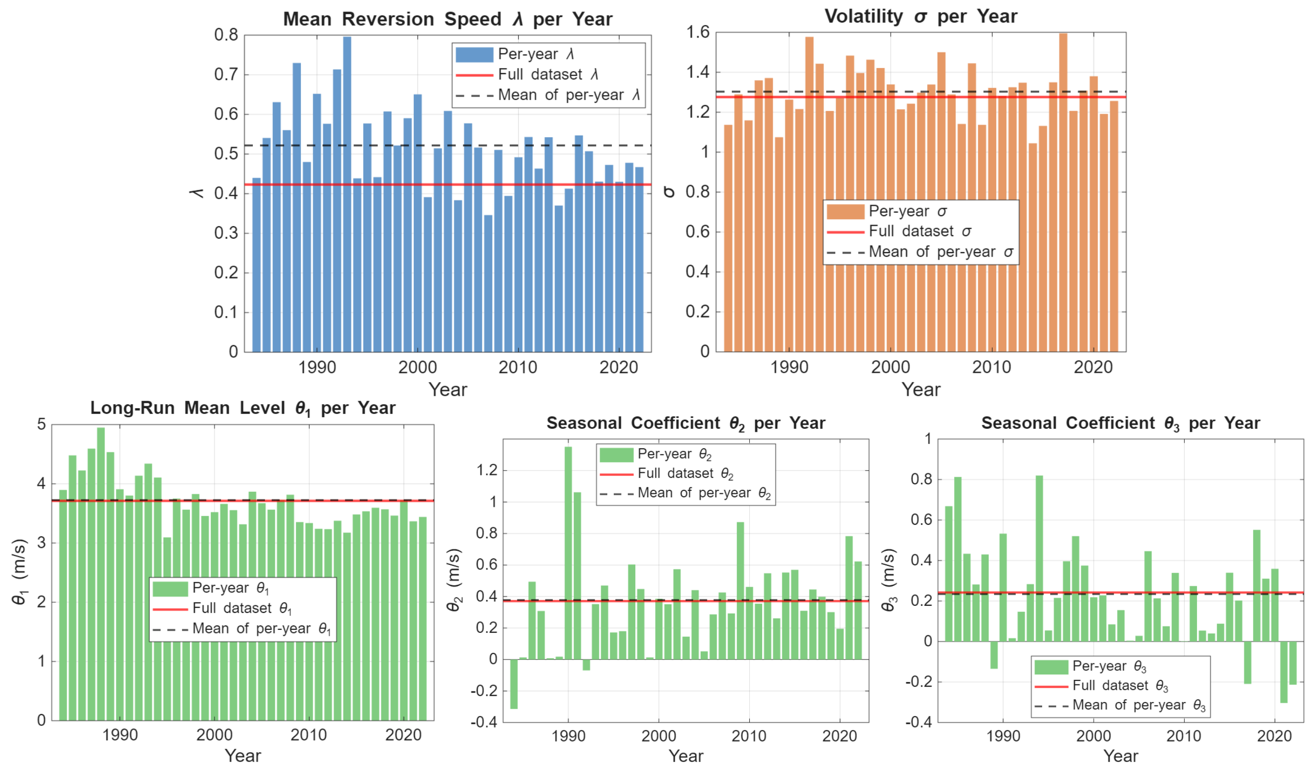

FIGURE 2. The bar plots of λ (top left), σ (top right), θ_1 (bottom left), θ_2 (bottom center), and θ_3 (bottom right) calculated values per year and over the entire dataset 2.

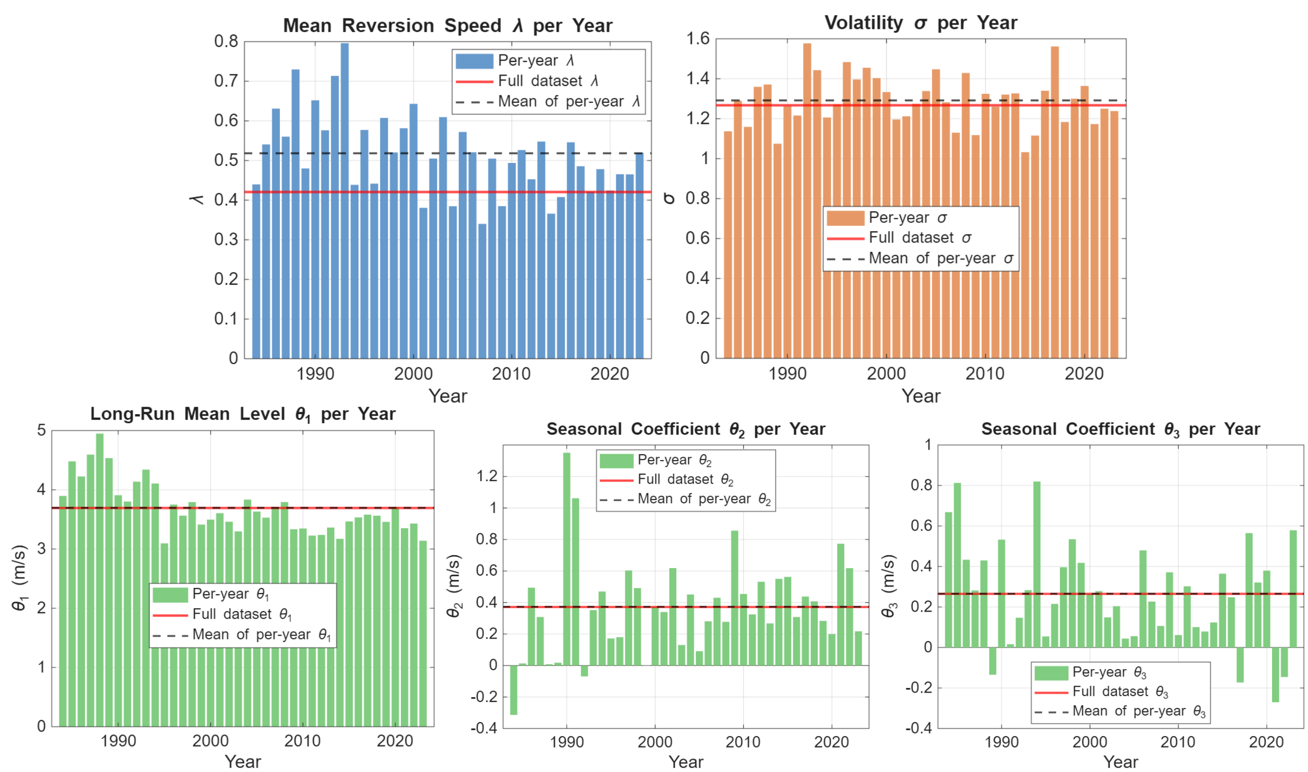

FIGURE 3. The bar plots of λ (top left), σ (top right), θ_1 (bottom left), θ_2 (bottom center), and θ_3 (bottom right) calculated values per year and over the entire dataset 3.

Since we're working with two different datasets that contain wind speed data from the same location, it's worth comparing the parameter value results from dataset 1 and dataset 2 to the parameter values estimated from a dataset that combines dataset 1 and 2.

4.6. **Combining Datasets 1 and 2.** In this section, datasets 1 and 2 are merged together by averaging the values where both are available and filling in the gaps with whichever dataset has data. Then, we will refer to this dataset as dataset 3, which contains $14,846$ observations from January 1st, 1984 to August 24, 2024 with the following mean and standard deviations:

$$\text{mean speed} = 3.6943 \text{ (m/s)}, \qquad \text{std speed } = 1.4191 \text{ (m/s)}.$$

In Figure 3, the bar graphs of each parameter value for each year are shown with the solid red line being the value calculated over the entire dataset and the dashed black lie is the mean of the per-year parameter value. For exact values, refer to Section 7 for the full table. For a full comparison, Figure 4 shows the AR(1) plots from each dataset 1, 2, and 3.

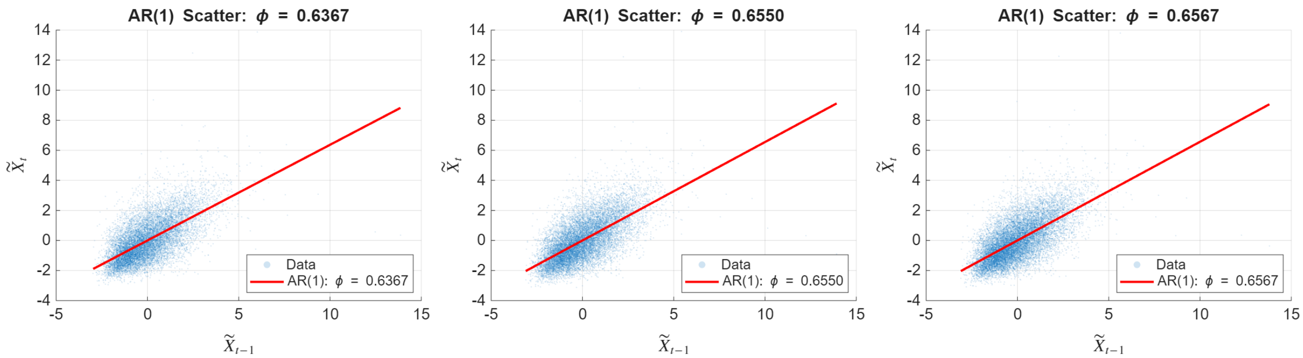

FIGURE 4. The AR(1) plots for dataset 1 (left), dataset 2 (center), and dataset 3 (right).

Concluding this section, we implemented a procedure using ordinary least squares and AR(1) regression on the residuals to compute the parameter values based off two version of the Miami International Airport wind speed time series. For the next section, the numerical results using the parameter values calculated from dataset 3 are shown.

5. Numerical Results

In this section, we implement the numerical schemes detailed in Section 3 and initial with the estimated parameters from Table 3, calculated by following the procedure detailed in Section 4.

6. Conclusion

7. Appendix

In Table 1, the parameters are estimated based off dataset 1's wind speed data for each year and over the entire 33-year period. Note, the table omits years with less than 300 observations. In Table 2, the calculated parameter values per year are listed, followed by the corresponding mean and standard deviation, then the parameter values calculated from the entire dataset. Note, the table omits years with less than 300 observations. Then, Table 3 lists the calculated parameter values per year, the corresponding mean and standard deviation, then the parameter value calculated over the entire dataset.

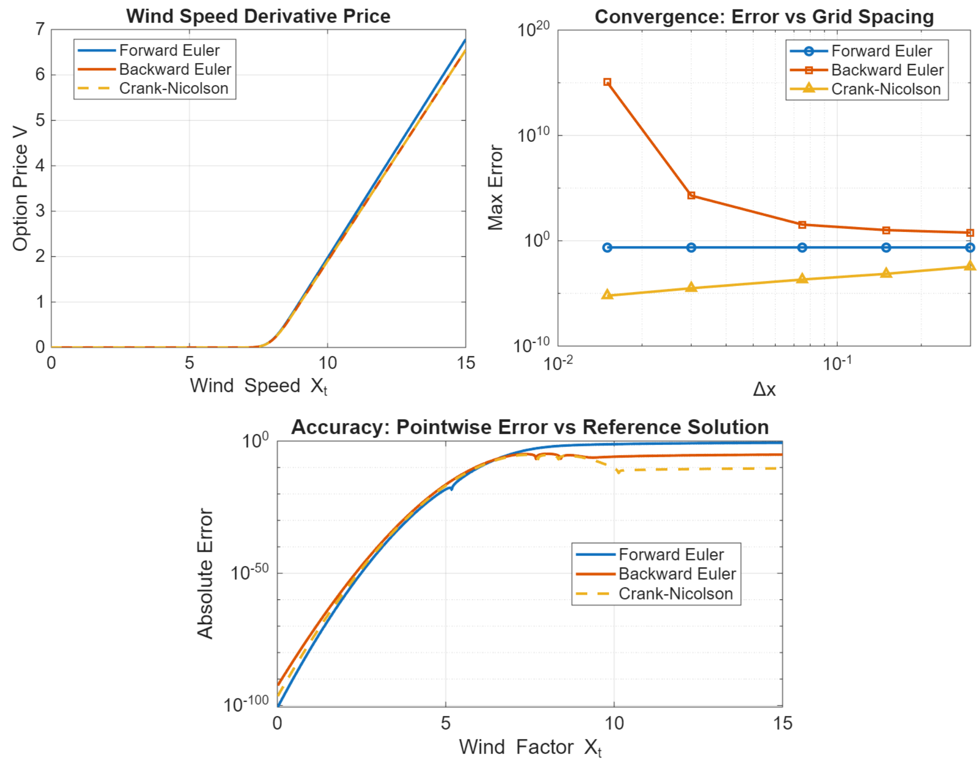

FIGURE 5. Caption

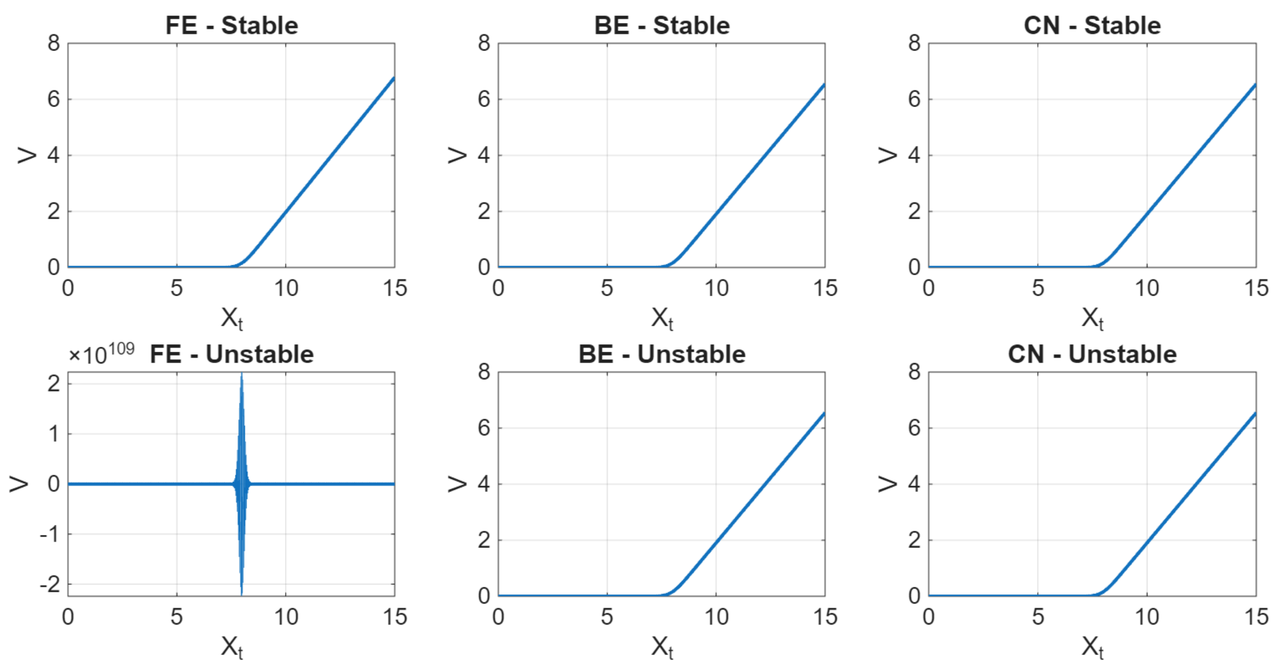

FIGURE 6. Caption

REFERENCES

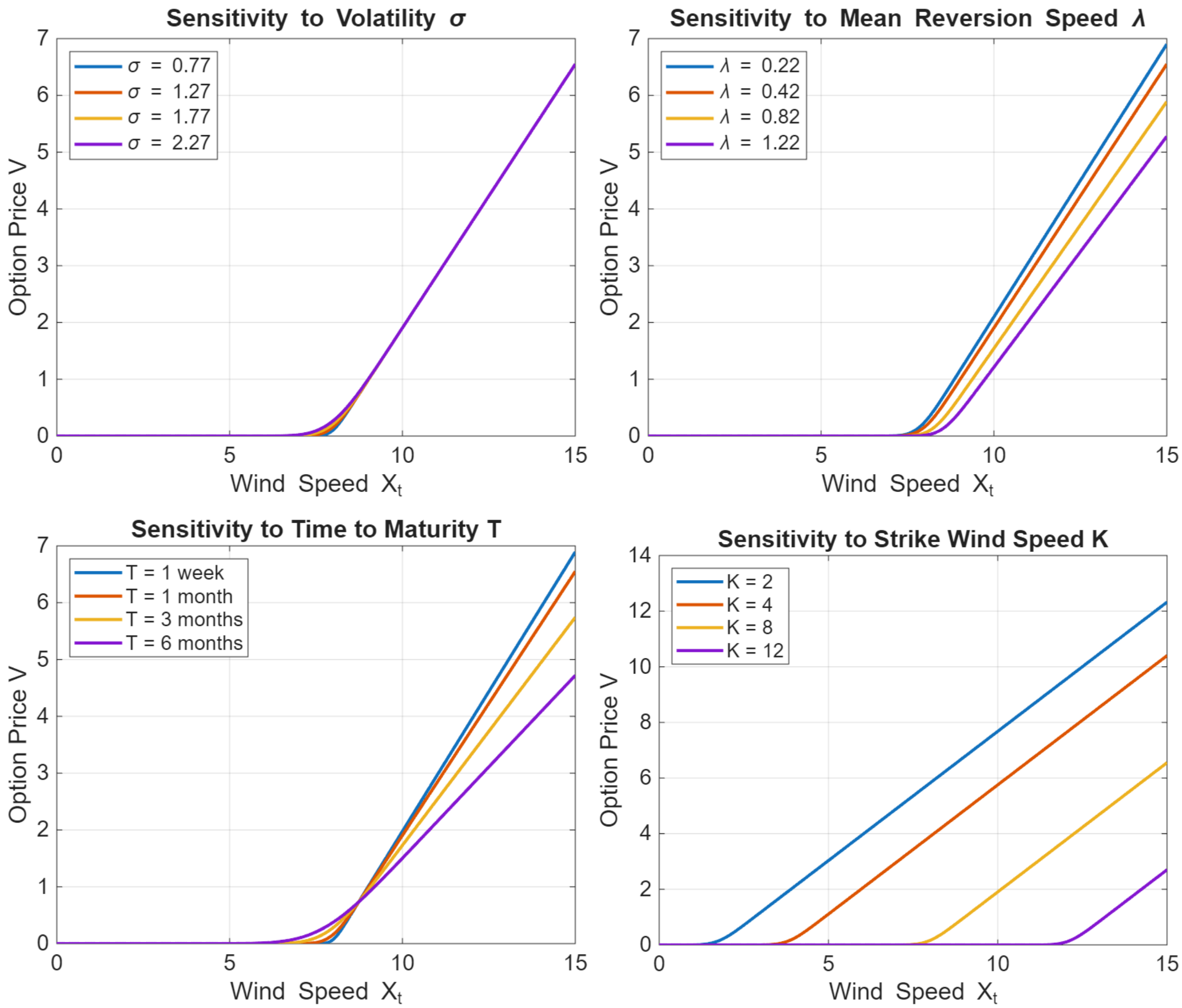

Figure 7

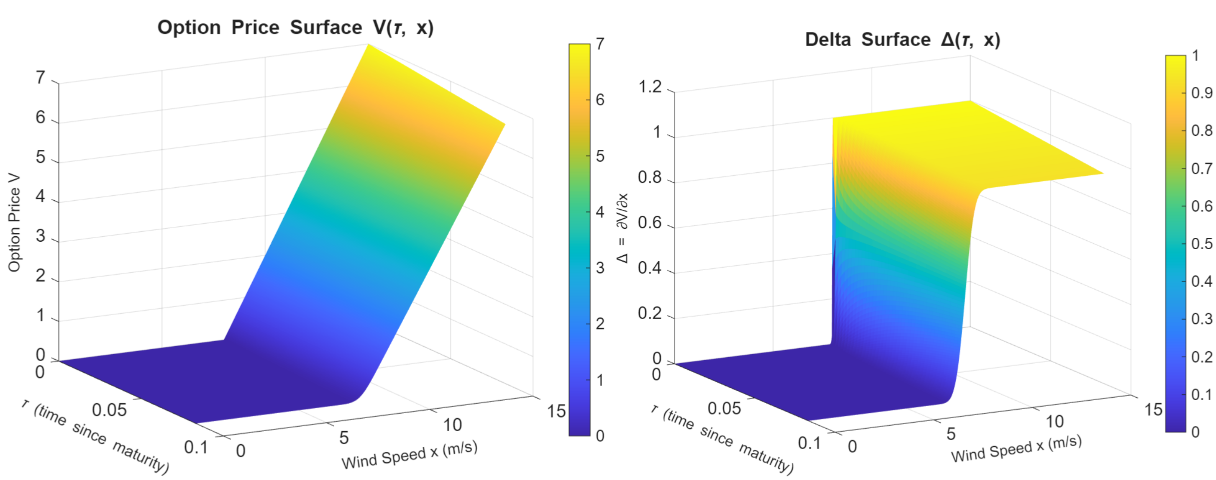

Figure 8. Caption

Parameters Estimated per Year from Dataset 1						
Year	λ	σ	θ_1	θ_2	θ_3	**# of Obs**
1990	0.6519	1.2631	3.9083	1.3508	0.5326	365
1991	0.5763	1.2160	3.8012	1.0620	0.0165	365
1992	0.7134	1.5773	4.1369	-0.0690	0.1471	366
1993	0.7959	1.4429	4.3401	0.3515	0.2833	365
1994	0.4386	1.2060	4.1055	0.4696	0.8198	365
1995	0.5771	1.2703	3.0950	0.1718	0.0546	365
1996	0.4417	1.4833	3.7496	0.1797	0.2154	366
1997	0.6074	1.3962	3.5628	0.6031	0.3963	365
1998	0.5215	1.4512	3.7203	0.5786	0.5630	365
1999	0.5707	1.3978	3.3264	-0.0221	0.5037	365
2000	0.6370	1.3452	3.4425	0.3526	0.3547	366
2001	0.3768	1.1982	3.5488	0.3286	0.3287	365
2002	0.5026	1.2081	3.3674	0.6634	0.2130	365
2003	0.6241	1.2658	3.2580	0.0999	0.3022	365
2004	0.3961	1.3669	3.7708	0.4705	0.1238	366
2005	0.5713	1.3945	3.5538	0.1682	0.1114	365
2006	0.5426	1.3000	3.4644	0.2683	0.5471	365
2007	0.3397	1.1407	3.6509	0.4404	0.2535	365
2008	0.5043	1.4322	3.7484	0.2458	0.1684	366
2009	0.3771	1.1137	3.2911	0.8236	0.4359	365
2010	0.5139	1.3756	3.3657	0.4400	0.1801	365
2011	0.5057	1.2516	3.1990	0.2649	0.3584	365
2012	0.4401	1.3450	3.2448	0.5000	0.1941	366
2013	0.5712	1.3234	3.3324	0.2796	0.1576	365
2014	0.3812	1.0652	3.1683	0.5452	0.1950	365
2015	0.4111	1.1210	3.4328	0.5486	0.4146	365
2016	0.5582	1.3626	3.5276	0.3049	0.3387	366
2017	0.4755	1.5676	3.5472	0.4255	-0.0992	364
2018	0.4321	1.1884	3.5381	0.4095	0.6176	365
2019	0.4984	1.3213	3.4480	0.2479	0.3430	365
2020	0.4277	1.3784	3.6920	0.2071	0.4137	366
2021	0.4513	1.1758	3.3205	0.7536	-0.2053	365
2022	0.4778	1.2864	3.4228	0.6024	-0.0487	365
2023	0.5379	1.2738	3.1172	0.2121	0.6021	365
Mean	0.5132	1.3090	3.5353	0.4200	0.2892	-
Std	0.1020	0.1243	0.2931	0.2868	0.2207	-
Parameters Estimated over the Entire Dataset 1						
-	0.4515	1.2948	3.5340	0.4195	0.2864	-

TABLE 1. The parameters estimated per year of wind speed observation data from year 1990 to 2023, followed by the mean and std per parameter over the 33-year period. The last row are the calculate values of each parameter over the entire dataset of observations over 33 years.

Parameters Estimated per Year from Dataset 2						
Year	λ	σ	θ_1	θ_2	θ_3	**# of Obs**
1984	0.4398	1.1366	3.8967	-0.3139	0.6682	366
1985	0.5406	1.2893	4.4792	0.0132	0.8126	365
1986	0.6308	1.1589	4.2253	0.4940	0.4332	365
1987	0.5602	1.3593	4.5940	0.3078	0.2818	365
1988	0.7298	1.3712	4.9488	0.0082	0.4300	366
1989	0.4799	1.0745	4.5350	0.0179	-0.1346	365
1990	0.6519	1.2631	3.9083	1.3508	0.5326	365
1991	0.5763	1.2160	3.8012	1.0620	0.0165	365
1992	0.7134	1.5773	4.1369	-0.0690	0.1471	366
1993	0.7959	1.4429	4.3401	0.3515	0.2833	365
1994	0.4386	1.2060	4.1055	0.4696	0.8198	365
1995	0.5771	1.2703	3.0950	0.1718	0.0546	365
1996	0.4417	1.4833	3.7496	0.1797	0.2154	366
1997	0.6074	1.3962	3.5628	0.6031	0.3963	365
1998	0.5216	1.4630	3.8270	0.4476	0.5203	365
1999	0.5904	1.4214	3.4569	0.0129	0.3751	365
2000	0.6505	1.3389	3.5213	0.3858	0.2183	366
2001	0.3913	1.2144	3.6612	0.3508	0.2280	365
2002	0.5144	1.2427	3.5546	0.5734	0.0845	365
2003	0.6087	1.2977	3.3168	0.1451	0.1550	365
2004	0.3836	1.3385	3.8667	0.4406	0.0040	366
2005	0.5775	1.5004	3.6721	0.0517	0.0281	365
2006	0.5163	1.2890	3.5626	0.2866	0.4460	365
2007	0.3461	1.1412	3.7283	0.4253	0.2130	365
2008	0.5104	1.4449	3.8155	0.2927	0.0753	366
2009	0.3944	1.1368	3.3542	0.8726	0.3391	365
2010	0.4918	1.3215	3.3369	0.4609	0.0021	365
2011	0.5430	1.2813	3.2406	0.3542	0.2740	365
2012	0.4632	1.3251	3.2346	0.5474	0.0541	366
2013	0.5424	1.3479	3.3764	0.2614	0.0395	365
2014	0.3701	1.0444	3.1751	0.5524	0.0881	365
2015	0.4127	1.1315	3.4821	0.5693	0.3395	365
2016	0.5470	1.3497	3.5360	0.3094	0.2024	366
2017	0.5070	1.5950	3.5954	0.4449	-0.2094	364
2018	0.4305	1.2066	3.5693	0.3984	0.5515	364
2019	0.4725	1.3078	3.4628	0.3005	0.3105	365
2020	0.4302	1.3799	3.7216	0.1954	0.3589	366
2021	0.4778	1.1914	3.3667	0.7832	-0.3037	365
2022	0.4670	1.2564	3.4401	0.6226	-0.2132	365
Mean	0.5216	1.3029	3.7244	0.3777	0.2343	-
Std	0.1028	0.1306	0.4260	0.3098	0.2597	-
Parameters Estimated over the Entire Dataset 2						
-	0.4231	1.2759	3.7118	0.3707	0.2419	-

Table 2. The parameters estimated per year of wind speed observation data from year 1984 to 2023, followed by the mean and std per parameter over the 40-year period. The last row are the calculate values of each parameter over the entire dataset of observations over 40 years.

Parameters Estimated per Year from Dataset 3						
Year	λ	σ	θ_1	θ_2	θ_3	**# of Obs**
1984	0.4398	1.1366	3.8967	-0.3139	0.6682	366
1985	0.5406	1.2893	4.4792	0.0132	0.8126	365
1986	0.6308	1.1589	4.2253	0.4940	0.4332	365
1987	0.5602	1.3593	4.5940	0.3078	0.2818	365
1988	0.7298	1.3712	4.9488	0.0082	0.4300	366
1989	0.4799	1.0745	4.5350	0.0179	-0.1346	365
1990	0.6519	1.2631	3.9083	1.3508	0.5326	365
1991	0.5763	1.2160	3.8012	1.0620	0.0165	365
1992	0.7134	1.5773	4.1369	-0.0690	0.1471	366
1993	0.7959	1.4429	4.3401	0.3515	0.2833	365
1994	0.4386	1.2060	4.1055	0.4696	0.8198	365
1995	0.5771	1.2703	3.0950	0.1718	0.0546	365
1996	0.4417	1.4833	3.7496	0.1797	0.2154	366
1997	0.6074	1.3962	3.5628	0.6031	0.3963	365
1998	0.5205	1.4549	3.7914	0.4913	0.5345	365
1999	0.5814	1.4034	3.4134	0.0013	0.4180	365
2000	0.6430	1.3333	3.4967	0.3718	0.2650	366
2001	0.3807	1.1958	3.6050	0.3397	0.2784	365
2002	0.5049	1.2118	3.4610	0.6184	0.1488	365
2003	0.6095	1.2745	3.2972	0.1301	0.2041	365
2004	0.3846	1.3382	3.8347	0.4506	0.0439	366
2005	0.5720	1.4475	3.6327	0.0905	0.0559	365
2006	0.5209	1.2820	3.5298	0.2805	0.4797	365
2007	0.3402	1.1300	3.7025	0.4303	0.2265	365
2008	0.5047	1.4288	3.7931	0.2770	0.1063	366
2009	0.3849	1.1178	3.3332	0.8563	0.3713	365
2010	0.4940	1.3243	3.3465	0.4540	0.0614	365
2011	0.5264	1.2601	3.2267	0.3244	0.3021	365
2012	0.4526	1.3209	3.2380	0.5316	0.1007	366
2013	0.5477	1.3267	3.3617	0.2675	0.0789	365
2014	0.3659	1.0326	3.1728	0.5500	0.1238	365
2015	0.4079	1.1157	3.4657	0.5624	0.3645	365
2016	0.5461	1.3397	3.5332	0.3079	0.2478	366
2017	0.4856	1.5618	3.5796	0.4378	-0.1728	364
2018	0.4217	1.1833	3.5631	0.4072	0.5648	365
2019	0.4782	1.3000	3.4578	0.2830	0.3211	365
2020	0.4244	1.3642	3.7110	0.1987	0.3797	366
2021	0.4653	1.1731	3.3517	0.7730	-0.2704	365
2022	0.4651	1.2502	3.4299	0.6178	-0.1456	365
2023	0.5196	1.2383	3.1390	0.2173	0.5790	365
Mean	0.5183	1.2913	3.6961	0.3729	0.2656	-
Std	0.1029	0.1279	0.4349	0.3065	0.2519	-
Parameters Estimated over the Entire Dataset 3						
-	0.4205	1.2673	3.6928	0.3713	0.2648	-

TABLE 3. The parameters estimated per year of wind speed observation data from dataset 3, followed by the mean and std per parameter over the 40-year period. The last row are the calculate values of each parameter over the entire dataset 3 of observations over 40 years.