Simultaneous optimization for wind derivatives based on prediction errors

Yuji Yamada

Abstract—Wind power energy has been paid much attention recently for various reasons, and the production of electricity with wind energy has been increasing rapidly for a few decades. One of the most difficult issues for using wind power in practice is that the power output largely depends on the wind condition, and as a result, the future output may be volatile or uncertain. Therefore, the prediction of power output in the future is considered important and is key to electric power generating industries making the wind power electricity market work properly. However, the use of predictions may cause other problems due to “prediction errors.”

In this work, we will propose a new type of weather derivatives based on the prediction errors for wind speeds, and estimate their hedge effect on wind power energy businesses. At first, we will investigate the correlation of prediction errors between the power output and the wind speed in a Japanese wind farm, which is a collection of wind turbines that generate electricity in the same location. Then we will develop a methodology that will optimally construct a wind derivative based on the prediction errors using nonparametric regressions. A simultaneous optimization technique of the loss and payoff functions for wind derivatives is demonstrated based on the empirical data.

Keywords: Wind power energy, Prediction errors, Weather derivatives, Minimum variance hedge, Non-parametric regression

I. INTRODUCTION

Predicting the future weather conditions is considered important in real businesses for many industries including electricity producers and suppliers, because their profit or loss is largely affected by the weather conditions. Under these circumstances, we may have a new risk when the prediction error exists. In this work, we will propose a new type of weather derivative (see, e.g., [3] for the introduction of weather derivatives) to effectively hedge the loss caused by prediction errors of power output for wind power electricity production.

Electricity companies must sell the output immediately because the electricity has to be consumed as soon as it is produced. Therefore, sales contracts need to be written in advance. However, in the case of electricity production using wind power energy, the power output largely depends on wind conditions, and as a result, tradable volume is uncertain. What we can do is to predict the future outputs and quote them in advance. But, this may cause another risk associated with prediction errors. The objective of this work is to hedge the loss using weather derivatives based on the wind speed.

The idea is to construct such a derivative based on prediction errors of wind speed. In contrast to the standard weather derivatives in which the underlying index is given by weather data only (such as temperature [1], [2], [5], [6], [7], [8]), the proposed weather derivative uses prediction data and the payoff depends on the difference between the actual data and the prediction data. For simplicity, we shall call this type of weather derivatives contract just “wind derivatives.”

Here we consider the prediction of the power output from a wind farm (WF), which is a collection of wind turbines that generate electricity in the same location. The power output is predicted using numerical weather prediction where a public weather forecasting company compute sophisticated values from Japan Meteorological Agency data. Because of this prediction mechanism, we have both the wind and power predictions data.

The value of electricity generated by wind power is normally considered to be low due to the uncertainty of the tradable volume. Here we assume that the electricity price without prediction is estimated to be 3 yen per 1 kWh. On the other hand, the value of the electricity would be estimated to be higher, if the tradable volume were quoted in advance by prediction, but the buyer has to guarantee the quoted volume or has to pay the penalty in case of shortages. Suppose that the value of electricity with prediction is given as 7 yen per 1 kWh and that the penalty of the shortage is 10 yen per 1 kWh. These assumptions are not so far from the current situation discussed in the prediction business. In this case, the loss function caused by prediction errors is depicted in Fig. 1, which shows the relation between the prediction error for the power output $P - \hat{P}$ (the actual power output minus its prediction) and the loss caused by the prediction error. Here we also consider an opportunity loss where the actual output is greater than the prediction.

One of the objectives of this work is, given the loss function, to find the optimal payoff structure of weather derivatives on prediction errors of wind speed using a non-parametric regression. To this end, we will first consider the following problems:

P1) Given the loss function and the payoff function of wind derivatives, find the optimal volume of wind derivative using linear regression.

P2) Given the loss function, find the optimal payoff function of wind derivatives using GAM.

We will investigate the hedge effect of wind derivatives and show that using wind derivatives on prediction error of
A. Generalized additive models

We will solve a non-parametric regression problem in this paper to find a (cubic) smoothing spline that minimizes the penalized residual sum of squares (PRSS) among all regression spline functions with two continuous derivatives. Let \( y_n \) and \( x_n \) be dependent and independent variables, respectively, and express \( y_n \) as

\[
y_n = h(x_n) + \varepsilon_n, \quad \text{Mean}(\varepsilon_n) = 0
\]

(1) using a function \( h(\cdot) \) and residuals \( \varepsilon_n \), where Mean(\( \cdot \)) is a sample mean. Here the function \( h(\cdot) \) is a (cubic) smoothing spline that minimizes the following penalized sum of squares (PRSS; see e.g., [4]),

\[
\text{PRSS} = \sum_{i=1}^{n} \{y_n - h(x_n)\}^2 + \lambda \int \{h''(x)\}^2 \, dx
\]

(2) among all functions \( h(\cdot) \) with two continuous derivatives. In (2), the first term measures closeness to the data while the second term penalizes curvature in the function. Note that, if \( \lambda = 0 \) and \( h(\cdot) \) is given by a polynomial function, the problem is reduced to the standard regression polynomial and is solved by the least squares method. It is shown that (2) has an explicit, unique minimizer, and that regression splines can be extended to the multivariable case with additive sums of smoothing splines, known as generalized additive models (GAMs) [4]. Note that \( \lambda \) may be found by using the so-called generalized cross validation criteria. Also note that GAMs can be computed using free software “R (http://cran.r-project.org),” and we will refer to the class of smoothing splines for non-parametric regression as GAMs in this paper.

In the following sections, we will apply GAMs to solve the problems 1–4 stated in Section I and estimate the hedge effect of wind derivatives. Before demonstrating our simulation results, we will briefly explain the prediction technique of wind conditions and the power output of a WF, and demonstrate individual data analysis.

B. Basic idea of power prediction and data set

The output from the WF is predicted based on the numerical weather prediction and the power generating properties for turbines. The numerical weather prediction consists of the following two steps:

- Japan Meteorological Agency announces the hourly data of regional spectral models for the next 51 hours twice a day (9am and 9pm).
- Using them as initial and boundary values, a public weather forecasting company computes more sophisticated values for the next day’s hourly data by 12pm.

In this paper, we use the prediction data obtained from the Local Circulation Assessment and Prediction System (LOCALS) developed by the ITOCHU Techno-Solutions Corporation for the wind speed and the power output of a wind farm in Japan. The data set is given as follows:\(^1\)

\begin{itemize}
  \item Realized and predicted values of total power output for the WF, and those of wind speed for the observation tower in the WF.
\end{itemize}

\begin{itemize}
  \item Data period: 2002–2003 (1 year), hourly data, everyday
  \item Total number of data: 8,000 for each variable excluding missing values
\end{itemize}

Let \( n = 1, \ldots, N \) be the time index and define the following variables:

\[
\begin{align*}
P_n & : \text{Total power output at time } n \\
P_n^* & : \text{Prediction of } P_n \\
W_n & : \text{Wind speed at time } n \\
W_n^* & : \text{Prediction of } W_n
\end{align*}
\]

We will examine the relationships between the above variables below: Fig. 2 shows a scatter diagram for the wind speed \( W_n \) and the power output \( P_n \), where the power output \( P_n \) is normalized so that its maximum equals 100. From Fig. 2, we can see that:

- The generator starts providing the power output when the wind speed exceeds around 2 [m/s].
- The power output increases with the wind speed between 5–15 [m/s].

Also note that, because each electricity generator is controlled so that the maximum output does not exceed a certain value, the total output is also bounded as shown in Fig. 2.

\(^1\)All the data used in this paper were provided by ITOCHU Techno-Solutions Corporation.
Fig. 2. Wind speed $W_n$ [m/s] vs. Power output $P_n$ [W]

Fig. 3 shows a partial residual plot for

$$P_n = a_p \hat{P}_n + b_p + \varepsilon_{p,n}, \quad n = 0, \ldots, N, \quad \text{Mean}(\varepsilon_{p,n}) = 0$$  \hspace{1cm} (3)

i.e., the scatter diagram of $(\hat{P}_n, P_n - b_p)$, where $a_p$ and $b_p$ are a regression coefficient and intercept, respectively, and $\varepsilon_{p,n}$ is a residual satisfying Mean($\varepsilon_{p,n}$) = 0, the partial regression line is depicted using a solid straight line shown in Fig. 3. In this case, the sample variance of residuals is given as

$$\text{Var}(\varepsilon_{p,n}) = 249.$$  \hspace{1cm} (4)

On the other hand, the regression spline $g(\cdot)$ to fit the same data of Fig. 3 is shown as a solid line in Fig. 4, where $g(\cdot)$ satisfies

$$P_n = g(\hat{P}_n) + \varepsilon_{p,n}.$$  \hspace{1cm} (5)

using GAM. In this case, the variance of the residuals is given as

$$\text{Var}(\varepsilon_{p,n}) = 239.$$  \hspace{1cm} (6)

Noting that the variance of the measured values is computed as “504,” we can say that the variance of the power output is reduced by 50% (from “504” to “249”) using the predicted value and the linear regression, and it is improved a little using GAM, i.e., from “249” to “239.”

Fig. 4. Spline regression function for the power output using GAM

Fig. 5. Predicted vs. Measured values for the wind speed

Also, we draw a partial residual plot for the wind speed $W_n$ with respect to the predicted value $\hat{W}_n$ as shown in Fig. 5, where the solid line is obtained from a linear regression for partial residuals. In this case, the variance of the residuals is given as “5.12.” The solid line in Fig. 6 refers to the regression spline function $f(\cdot)$ satisfying

$$W_n = f(\hat{W}_n) + \varepsilon_{w,n}, \quad n = 0, \ldots, N$$  \hspace{1cm} (7)

using GAM. Note that the variance of residuals in this case is given as “4.95,” whereas the variance of the measured value of the wind speed is “11.0.” Similar to the power output case, we can say that the variance of the wind speed is reduced to less than half (from “11.0” to “5.12”) using the predicted value and the linear regression, and it is improved a little using GAM, i.e., “5.12” to “4.95.”

Throughout this paper, we define the prediction errors of the power output and the wind speed as the ones given by GAMs, i.e., $\varepsilon_{p,n}$ in (5) and $\varepsilon_{w,n}$ in (7), respectively.

Fig. 6. Spline regression function for the wind speed

III. OPTIMIZATION OF DERIVATIVE CONTRACTS

In this section, we will formulate the first two optimization problems, i.e., P1) and P2), stated at the end of Section I. Let $\varepsilon_{p,n}$ and $\varepsilon_{w,n}$ be given by (5) and (7), respectively. Assume that there is an investor whose loss function with respect to $\varepsilon_{p,n}$ is given by $\phi(\varepsilon_{p,n})$. For instance, the loss function may be given as the one shown in Fig. 1 for a WF owner. Without loss of generality, we will assume that

$$\text{Mean}(\phi(\varepsilon_{p,n})) = 0.$$  \hspace{1cm} (8)
A. Minimum variance hedge

Consider a situation in which an investor with a loss function $\phi(\varepsilon_{p,n})$ would like to compensate their loss on $\varepsilon_{p,n}$ using a wind derivative on $\varepsilon_{w,n}$. Let $\psi(\varepsilon_{w,n})$ be a payoff of the derivative contract at time $n$, where $\psi(\cdot)$ is a payoff function. Assume that the contracts are carried out in advance without any cost and that $\psi(\varepsilon_{w,n})$ satisfies the following condition:

$$\text{Mean}(\psi(\varepsilon_{w,n})) = 0.$$  \hspace{1cm} (9)

Note that condition (9) indicates that the physical measure provides a risk-neutral measure, and that, in the case of simple forward contracts, $\psi(\varepsilon_{w,n})$ may be given as a linear function, e.g.,

$$\psi(\varepsilon_{w,n}) = \varepsilon_{w,n}.  \hspace{1cm} (10)$$

With these definitions, the first optimization problem is formulated as follows:

**Contract volume optimization problem:**

$$\min_{\Delta \in \mathbb{R}} \text{Var}(\phi(\varepsilon_{p,n}) + \Delta \psi(\varepsilon_{w,n})) \hspace{1cm} (11)$$

where $\text{Var}(\cdot)$ is a sample variance. Note that the above problem is also known as the “minimum variance hedging problem.” For this problem, the optimal volume $\Delta^*$ may be computed analytically as

$$\Delta^* = -\frac{\text{Cov}(\phi(\varepsilon_{p,n}), \psi(\varepsilon_{w,n}))}{\text{Var}(\psi(\varepsilon_{w,n}))}. \hspace{1cm} (12)$$

To estimate the hedge effect, we define the variance reduction rate (VRR) as follows:

$$\text{VRR} := \frac{\text{Var}(\phi(\varepsilon_{p,n}) + \Delta^* \psi(\varepsilon_{w,n}))}{\text{Var}(\phi(\varepsilon_{p,n}))}. \hspace{1cm} (13)$$

Because the minimum variance can be computed as

$$\text{Var}(\phi(\varepsilon_{p,n}) + \Delta^* \psi(\varepsilon_{w,n})) = \text{Var}(\phi(\varepsilon_{p,n})) \left(1 - |\text{Corr}(\phi(\varepsilon_{p,n}), \psi(\varepsilon_{w,n}))|^2\right),$$

we obtain

$$\text{VRR} = 1 - |\text{Corr}(\phi(\varepsilon_{p,n}), \psi(\varepsilon_{w,n}))|^2. \hspace{1cm} (14)$$

Note that VRR satisfies

$$0 \leq \text{VRR} \leq 1 \hspace{1cm} (15)$$

and that a smaller VRR provides a better hedge effect in terms of minimum variance.

B. Optimization with payoff functions

In the minimum variance hedge, we computed the optimal volume for a given payoff function of a wind derivative contract. More generally, we can optimize the payoff function as well using a similar manner to GAM as follows: Consider the minimization problem of PRSS in (2) for a given $\lambda$. Instead of writing the problem as an unconstrained optimization problem, we can reformulate it as an optimization problem constrained on $h(\cdot)$ as follows:

$$\min_{h(\cdot)} \sum_{i=1}^{n} (y_n - h(x_n))^2 \hspace{1cm} (17)$$

s.t. $\int \{h''(x)\}^2 \, dx \leq \alpha$

where $\alpha$ is a given parameter. Note that the minimization problem in (17) is equivalent to

$$\min_{h(\cdot), \lambda} \left\{ \sum_{i=1}^{n} (y_n - h(x_n))^2 + \lambda \left( \int \{h''(x)\}^2 \, dx - \alpha \right) \right\} \hspace{1cm} (18)$$

using a Lagrange multiplier $\lambda > 0$ and that fixing $\lambda$ in (2) corresponds to fixing $\alpha$ in (18). Therefore, the non-parametric regression problem using GAM may be recast as a minimization problem of the sample variance with a smooth constraint. Now, we will formulate the second optimization problem below:

**Payoff function optimization problem:**

$$\min_{\psi(\cdot)} \text{Var}(\phi(\varepsilon_{p,n}) + \psi(\varepsilon_{w,n})) \hspace{1cm} (19)$$

s.t. $\int \{\psi''(x)\}^2 \, dx \leq \alpha$

The minimization problem in (19) may be recast as (17) by taking $y_n = \phi(\varepsilon_{p,n}), x_n = \varepsilon_{w,n}$, and $h(\cdot) = \psi(\cdot)$, and therefore, can be solved by applying GAM. Let $\psi''(\cdot)$ be the optimal payoff function. Then VRR may be defined as

$$\text{VRR} := \frac{\text{Var}(\phi(\varepsilon_{p,n}) + \psi''(\varepsilon_{w,n}))}{\text{Var}(\phi(\varepsilon_{p,n}))}. \hspace{1cm} (20)$$

In this case, although condition (15) does not hold exactly due to the smoothing condition in (19), we can still approximate VRR as

$$\text{VRR} \approx 1 - |\text{Corr}(\phi(\varepsilon_{p,n}), \psi''(\varepsilon_{w,n}))|^2. \hspace{1cm} (21)$$

with high accuracy in general.

IV. CONSTRUCTION OF WIND DERIVATIVES AND THEIR HEDGE EFFECT

In this section, we will construct wind derivatives and demonstrate their hedge effect on wind power businesses.

At first, we solve the minimum variance hedging problem for the simplest case where the loss and the payoff functions are both linear. Let

$$\phi(\varepsilon_{p,n}) = \varepsilon_{p,n}, \quad \psi(\varepsilon_{w,n}) = \varepsilon_{w,n} \hspace{1cm} (22)$$

Fig. 6. Spline regression function for the wind speed using GAM
without loss of generality. In this case, the problem is reduced to solving a linear regression for the following regression function:

\[ \varepsilon_{p,n} = a_w \varepsilon_{w,n} + \eta_{w,n} \]  
(23)

where \( \eta_{w,n} \) is a residual. Since the linear regression computes \( a_w \) that minimizes variance of \( \eta_{w,n} = \varepsilon_{p,n} - a_w \varepsilon_{w,n} \), the regression coefficient provides the optimal volume as

\[ \Delta^* = -a_w \]  
(24)

in (11) under condition (22), where

\[ a_w = \frac{\text{Cov}(\varepsilon_{p,n}, \varepsilon_{w,n})}{\text{Var}(\varepsilon_{w,n})} \]  
(25)

Fig. 7 shows a scatter plot of \( \varepsilon_{w,n} \) vs. \( \varepsilon_{p,n} \) with a linear regression line. The correlation is computed as

\[ \text{Corr}(\varepsilon_{p,n}, \varepsilon_{w,n}) \approx 0.70. \]  
(26)

and VRR as

\[ \text{VRR} = 1 - \text{Corr}(\varepsilon_{p,n}, \varepsilon_{w,n})^2 \approx 0.51. \]  
(27)

We see that the prediction errors of the wind speed and the power output, \( \varepsilon_{w,n} \) and \( \varepsilon_{p,n} \), are highly correlated and that the variance is reduced by 51\% using the wind derivative in the case where the loss and the payoff functions are both linear.

Next, we will consider the case in which the loss function \( \phi(\cdot) \) is given as shown in Fig. 1 with zero mean constraint (8), i.e.,

\[ \phi(\varepsilon_{p,n}) = 4|\varepsilon_{p,n}|^+ + 10|\varepsilon_{p,n}|^- - c \]  
(28)

where

\[ c := \text{Mean}(4|\varepsilon_{p,n}|^+ + 10|\varepsilon_{p,n}|^-). \]

and \(|\cdot|^+\) and \(|\cdot|^-\) are defined as

\[ |x|^+ := \max(x, 0), \quad |x|^- := \min(x, 0) \]

for \( x \in \mathbb{R} \). The solid line in Fig. 8 shows the optimal payoff function to solve the problem in (19). In this case, VRR in (20) is computed as

\[ \text{VRR} = 0.5461946 \cdots \]  
(29)

whereas the right hand side of (21) is

\[ 1 - [\text{Corr}(\phi(\varepsilon_{p,n}), \psi(\varepsilon_{w,n}))]^2 = 0.5461927 \cdots. \]  
(30)

From this example, we see that VRR can be approximated as in (21) with high accuracy.

**V. OPTIMIZATION WITH LOSS FUNCTIONS**

**A. Optimal loss function**

Next, we will consider a case in which a payoff function of wind derivative is given but we would like to find a loss function that is desirable for using the wind derivative, i.e., in a case where there already exists a standardized derivative contract with a certain payoff function, but there is some room for improvement on the loss function, e.g., for a WF owner. We assume that possible losses on \( \varepsilon_{p,n}, \phi(\varepsilon_{p,n}) \), has the same mean and variance, i.e., \( \phi(\varepsilon_{p,n}) \) satisfies

\[ \text{Mean}(\phi(\varepsilon_{p,n})) = 0, \]

\[ \text{Var}(\phi(\varepsilon_{p,n})) = c. \]  
(31)

We will compute an optimal loss function satisfying (31).

The optimization problem is formulated as follows:

**Loss function optimization problem:**

\[ \min_{\phi(\cdot)} \quad \text{Var}(\phi(\varepsilon_{p,n}) + \psi(\varepsilon_{w,n})) \]

\[ \text{s.t.} \quad \int \{\phi''(x)\}^2 \, dx \leq \alpha \]  
(32)

Let \( \hat{\phi}(\cdot) \) be the optimizer of problem (32), which can be computed by applying GAM. By normalizing \( \hat{\phi}(\cdot) \) to satisfy (31), we obtain the optimal loss function \( \phi^*(\cdot) \) as follows:

\[ \phi^*(\cdot) = \frac{c}{\text{Var}(\hat{\phi}(\varepsilon_{p,n}))} \hat{\phi}(\cdot) \]  
(33)

Note that the optimal volume of wind derivative with the payoff function \( \psi(\cdot) \) will be found by solving the standard minimum variance hedging problem as in Subsection III-A, and VRR may be computed as

\[ \text{VRR} = 1 - [\text{Corr}(\phi^*(\varepsilon_{p,n}), \psi(\varepsilon_{w,n}))]^2 \]  
(34)
B. Simultaneous optimization

It may be interesting to consider a simultaneous optimization of \( \phi(\varepsilon_{p,n}) \) and \( \psi(\varepsilon_{w,n}) \). Recall that VRR can be computed using the correlation between the payoff function and the loss function as

\[
1 - |\text{Corr}(\phi(\varepsilon_{p,n}), \psi(\varepsilon_{w,n}))|^2.
\]

Since the larger correlation the smaller VRR, the minimization of VRR boils down to the maximization of correlation between \( \phi(\varepsilon_{p,n}) \) and \( \psi(\varepsilon_{w,n}) \). Therefore, the simultaneous optimization of the payoff and the loss functions is formulated as follows:

**Simultaneous optimization problem:**

\[
\begin{align*}
\max_{\phi(\cdot), \psi(\cdot)} & \quad \text{Corr}(\phi(\varepsilon_{p,n}), \psi(\varepsilon_{w,n})) \\
\text{s.t.} & \quad \int \{ \phi''(x) \}^2 dx \leq \alpha\phi \\
& \quad \int \{ \psi''(x) \}^2 dx \leq \alpha\psi
\end{align*}
\]

(35)

The simultaneous optimization problem may be solved using an iterative algorithm by applying GAM with fixed \( \phi(\cdot) \) or \( \psi(\cdot) \) at each step. The following is the iterative algorithm:

**Iterative algorithm:**

1) Given \( \phi(\cdot) \), find \( \psi(\cdot) \) to solve the payoff function optimization problem. Let \( \psi^*(\cdot) \) be the optimal function, and let \( \psi(\cdot) = \psi^*(\cdot) \).
2) Given \( \psi(\cdot) \), find \( \phi(\cdot) \) to solve the loss function optimization problem. Let \( \phi^*(\cdot) \) be the optimal loss function and let \( \phi(\cdot) = \phi^*(\cdot) \).
3) Repeat Steps 2 and 3 until the correlation in (35) converges.

Note that the optimal loss function obtained from the above iterative algorithm satisfies (31) and that we can consider additional constraints to take more realistic situations into account for the loss and payoff functions.

C. Illustrative example

Here we will provide an illustrative example of solving P3) to compute an optimal loss function. Note that an example of P4) for simultaneous optimization is under preparation and will be given in the final manuscript.

Since the linear correlation between \( \varepsilon_{p,n} \) and \( \varepsilon_{w,n} \) is high in this example, it would be more interesting to consider the case where a payoff function is non-linear with respect to \( \varepsilon_{w,n} \). Therefore, we assume that there already exists a derivative contract with the payoff being proportional to the size of the wind speed prediction error \( |\varepsilon_{w,n}| \). Noting that \( \psi(\varepsilon_{w,n}) \) satisfies (9), such a payoff function may be given as

\[
\psi(\varepsilon_{w,n}) = |\varepsilon_{w,n}| - \text{Mean}(|\varepsilon_{w,n}|),
\]

(36)

Fig. 9 shows the payoff function with respect to \( \varepsilon_{w,n} \) given in (36).

VI. CONCLUDING REMARKS

In this work, we have proposed a new type of weather derivatives based on the prediction errors for wind speeds and estimated their hedge effect on wind power energy businesses. At first, we explained some properties of the loss for a WF caused by prediction errors of the power output, and characterized it using a loss function on the error. Then we introduced a non-parametric regression technique based on GAM and formulated an optimization problem to find the optimal payoff structure of weather derivatives on prediction errors of wind speed for the given loss function. Then we formulated and solved an optimization problem of the loss function for a given payoff function of the derivative contract on the wind speed prediction error using GAM. A simultaneous optimization technique of the loss and payoff functions for wind derivatives was also demonstrated.
REFERENCES