energies

MDPI

Article

Customized yet Standardized Temperature Derivatives: A Non-Parametric Approach with Suitable Basis Selection for Ensuring Robustness

Takuji Matsumoto [1] **and Yuji Yamada** [2,*]

[1] Socio-Economic Research Center, Central Research Institute of Electric Power Industry, Tokyo 100-8126, Japan; matsumoto3832@criepi.denken.or.jp
[2] Faculty of Business Sciences, University of Tsukuba, Tokyo 112-0012, Japan
* Correspondence: yuji@gssm.otsuka.tsukuba.ac.jp

Abstract: Previous studies have demonstrated that non-parametric hedging models using temperature derivatives are highly effective in hedging profit/loss fluctuation risks for electric utilities. Aiming for the practical applications of these methods, this study performs extensive empirical analyses and makes methodological customizations. First, we consider three types of electric utilities being exposed to risks of "demand", "price", and their "product (multiplication)", and examine the design of an appropriate derivative for each utility. Our empirical results show that non-parametrically priced derivatives can maximize the hedge effect when a hedger bears a "price risk" with high nonlinearity to temperature. In contrast, standard derivatives are more useful for utilities with only "demand risk" in having a comparable hedge effect and in being liquidly traded. In addition, the squared prediction error derivative on temperature has a significant hedge effect on both price and product risks as well as a certain effect on demand risk, which illustrates its potential as a new standard derivative. Furthermore, spline basis selection, which may be overlooked by modeling practitioners, improves hedge effects significantly, especially when the model has strong nonlinearities. Surprisingly, the hedge effect of temperature derivatives in previous studies is improved by 13–53% by using an appropriate new basis.

Keywords: electricity markets; non-parametric regression; minimum variance hedge; spline basis functions; cyclic cubic spline; weather derivatives

Citation: Matsumoto, T.; Yamada, Y. Customized yet Standardized Temperature Derivatives: A Non-Parametric Approach with Suitable Basis Selection for Ensuring Robustness. *Energies* **2021**, *14*, 3351. https://doi.org/10.3390/en14113351

Academic Editor: François Vallée

Received: 26 April 2021
Accepted: 3 June 2021
Published: 7 June 2021

1. Introduction

Electric utilities are generally exposed to the risk of daily fluctuations in price and demand, and constructing an efficient hedging methodology is an extremely important management issue. To this end, "electricity derivatives" may be introduced to prevent price fluctuations in electricity businesses. However, there is a potential problem that electricity derivatives may not be effective for "volume" (demand) risks. Moreover, for "price" risks as well, electricity derivatives may be unavailable in some markets, or their efficient use may be impossible because of low liquidity (especially for electricity derivatives with short-time granularity). In response to this awareness of issues, some studies have demonstrated the effectiveness of using "weather derivatives" instead of electric power derivatives. As examples of previous studies that verified the hedge effect of weather derivatives, Lee and Oren [1,2] discussed the effect of introducing standard weather derivatives into the hedging portfolio using equilibrium pricing models; however, their studies focused on the theory of suitable pricing instead of the empirical evaluation. Bhattacharya et al. [3] proposed the optimal trading strategy for standard derivatives based on heating degree days (HDD) and cooling degree days (CDD) using a data-driven approach. Their study empirically examined hedge effects, but as it optimized relatively simple two-dimensional vectors for the hedge weight of two different derivatives, the inherent nonlinear relationship

between price/demand and temperature was not necessarily incorporated. (Note that for existing standard weather derivatives discussed by these studies, many methods have been proposed in the context of "pricing" [4–7]).

To fill the research area of weather derivatives being uncovered in these past studies, non-parametric optimal hedging techniques for finding arbitrarily derivative payoff functions have been proposed in some studies (e.g., pricing method for derivatives on monthly average temperature) [8]; hedging for loss of power prediction errors for either wind [9,10] or solar power [11]; and simultaneous hedging method for electricity price and volume risks [12–14]). Of these, the most recent study [14] demonstrated that portfolios of weather derivatives may be constructed by applying generalized additive models (GAMs [15]), which provide a significantly high hedge effect for the fluctuation risks in electricity sales profit/loss defined by the "product of price and demand". The profit/loss to be hedged, which was defined as such, corresponds to the fluctuation risk of the procurement costs of "aggregators" who are procuring variable demand from the wholesale market at variable prices (if they re-sell the procured electricity to consumers at a fixed price, the hedged target corresponds to the fluctuation risk of the excess profits). However, some electric utilities are exposed only to price risk, while some are exposed only to volume risk; hence, it is necessary to pay particular attention to the fact that practically, each electricity business has different exposed risks. For instance, the electricity sales revenue of IPPs (independent power producers) that sell fixed power outputs generated by "base-load power plants" (i.e., coal-fired or nuclear power generation operated at rated output) at the wholesale market price bears only price risk. Similarly, retailers who supply power at a fixed price and volume to consumers (e.g., large factories) under a "base-load contract" and procure that volume from the wholesale market at a variable price also bear only price risks. On the contrary, electric utilities whose price risks are hedged by forward contracts want to hedge only volume risks. Especially in immature markets such as Japan, fixed-price bilateral contracts are widely concluded, wherein the daily supply volume can be flexibly changed, among retailers who have newly entered the market [16]; retailers who have such contracts are completely exposed only to volume risk. Hence, different types of power utilities have special needs to hedge individual risks, namely, for volume only or price only, and the development of a more customized hedging method proposed by [14] is still an open question. Therefore, the main purpose of this study is to conduct extensive empirical analysis for different types of business risks described above and to demonstrate practical applicability of the derivative models by using the empirical data from the PJM market, which is the world's largest regional transmission organization (RTO).

In order to deal with these different types of business risks, we explore how appropriate weather derivative product can be designed for each business risk. Our previous study [14] has provided the following two approaches using non-parametric hedging models to minimize fluctuations in daily revenues (cash flow) in terms of product design: (i) apply standard derivatives of weather values in which the number of contracts is optimized, and (ii) synthesize optimal derivative contracts using arbitrary payoff functions of weather values given the profit/loss structure of a hedger. There is a trade-off between the above two approaches, wherein case (i) has the advantage that the "customized yet standardized" derivatives can be traded liquidly among multiple players; while case (ii) uses a "made-to-order" derivative for each hedger, so it cannot be liquidly traded, but the hedge effect might be enhanced compared with case (i). In fact, the study [14] has demonstrated that case (ii) generally has a higher hedge effect than case (i) for the fluctuation risk defined by the "product" of price and demand. However, if the hedged target contains either only demand or price risk, the nonlinear effect of temperature or other weather indices is supposedly weakened; therefore, rather than designing a completely made-to-order derivative, as in previous studies [14], using "customized yet standardized" derivatives are expected to have the advantages of having sufficient hedging effects (or maybe comparable to the "made-to-order" type), as well as high versatility in that they can be traded among

multiple players. Note that the detailed transaction flow of "customized yet standardized" derivatives is found in Appendix A.

Furthermore, to ensure the out-of-sample hedging effect while using such standardized derivatives, methodological ingenuities to ensure robustness are considered very useful. Since non-parametric regressions with GAMs can express nonlinearities with more flexibility than parametric methods, we need to be aware of possible "overtraining"; therefore, when constructing various nonlinear models, it is necessary to estimate a function that expresses the nonlinearities "appropriately." For estimating nonlinear trends (i.e., payoffs or contract volumes) in hedging models, the R package "mgcv" [17] is very useful for practitioners in terms of its implementability and interpretability, as detailed in [18]. In "mgcv," various spline basis functions and smoothing methods are implemented, wherein the popular "cubic spline" (see, e.g., [19]) and "thin plate spline" [20] are set as the default basis (detailed in Section 3.3.1). However, the available basis functions include other advanced types, which may be overlooked by practitioners, such as "P-spline" [21], which has the advantage of "avoiding overtraining" [22], and "cyclic cubic spline" [23], which can robustly model periodic trends. Hence, if these bases are used instead, it is expected that robustness will be ensured, and the extrapolated hedging effects will be enhanced. In fact, our empirical result demonstrates that this hypothesis is correct and reveals that the hedge effect of temperature derivatives in previous studies is surprisingly improved by 13% to 53% by using an appropriate new basis function.

As described above, this study explores the important issues for decision-makers in the derivative contract practice, such as (i) what is an appropriate hedge product design for different business risks? and (ii) what is the appropriate spline basis function to enhance hedge effects? Then, by clarifying the interesting empirical results along with the theoretical interpretations, useful suggestions for practical application are provided.

This paper is organized as follows: Section 2 provides an overview of the background data of the PJM market, especially focusing on the nonlinearity of the data; Section 3 outlines the techniques used in this study; Section 4 formulates a specific hedging model treated in this paper. Section 5 examines the hedge effects of derivatives using empirical data. We also add considerations in the context of comparing business risk models or choice of bases in this section. Finally, Section 6 concludes the paper.

2. Overview of Background Data

This section provides an overview of the price and demand data in the PJM market, focusing on the "nonlinearities" that exist between data. First, Figure 1 illustrates a plot of demand in the PJM area with respect to the minimum and maximum temperatures. Both have a downwardly convex U-shaped curve, but they are relatively sensitive, especially when the temperature rises (i.e., in the summer). This figure suggests the existence of a strong nonlinear relationship between temperature and demand, but it is the story of looking at all the observation samples for one year. Considering that the temperature on each date fluctuates around the "climatological normal value," it is assumed that the fluctuation range of a certain date may be about 20 °F (11.1 °C) at the largest. This plot also indicates that the nonlinearity of the demand to temperature seems to be not very large, as long as temperature fluctuates within such range on the specific date.

Next, regarding the nonlinearity of price to demand, we overview the typical PJM "generation stack," shown in Figure 2. The generation stack is a curve wherein the marginal costs of supply capacity are arranged in ascending order, and it means the "supply curve." In other words, the generation stack corresponds to the plot of market price to demand (especially when assuming that the price elasticity of demand is 0, i.e., the demand curve is a vertical line). In practice, the generation stack can be frequently changed in the long- or short-term because of the termination/suspension of thermal power generation and the dependence of renewable power output on weather conditions; however, the point is that it is generally curved like a convex hockey stick. This strong nonlinearity of marginal cost curves is consistent with the fact that extreme price spikes occur when demand exceeds a

certain level in winter or summer. Note that although the generation stack implies that price is nonlinear to demand, because there is a strong correlation between temperature and demand, as seen in Figure 1, price is also inferred to be nonlinear to temperature.

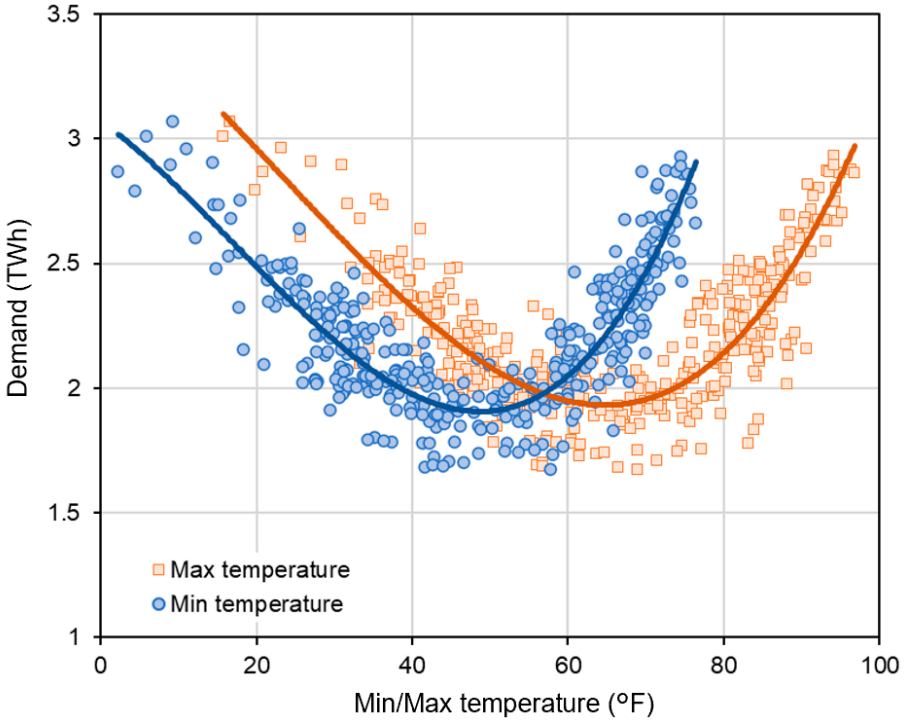

Figure 1. Relationship between temperature and electricity demand in PJM (2018).

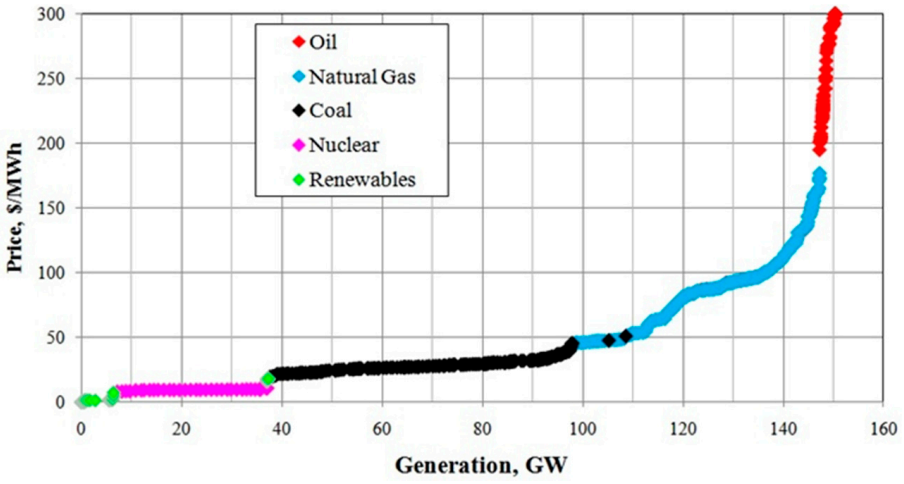

Figure 2. Typical PJM generation stack. Source: The Pennsylvania State University (https://www.e-education.psu.edu/ebf200/node/151) (accessed on 11 April 2021).

It should also be noted that the PJM price has a strong correlation with the Henry Hub (HH) natural gas price (see "Figure 5" in [14]). This corresponds to the fact that the natural gas part (blue line in Figure 2) of the generation stack shifts vertically because of fluctuations in the HH price. Because the demand curve and this supply curve intersect at the natural gas part in many time zones, electricity price may well be linked to the HH price as well. Additionally, considering that natural gas-fired power generation has increased significantly in recent years in the PJM area to replace coal-fired power generation, we will construct a hedge model that incorporates both the annual changes and the seasonal changes in the sensitivity of the HH price to the PJM price. Note that to target

the PJM market, this paper treats HH price as the representative fuel price that explains the electricity spot price (i.e., we use HH futures as hedge products for the fuel-linked price risk). However, when constructing a hedging model for another market, it is necessary to properly select fuel futures that have a strong correlation with the electricity price in that country/area. For instance, WTI crude oil futures have been reported to be effective in hedging Japanese electricity market prices [14].

3. Minimum Variance Hedging Problem

In this study, we consider the problem of minimizing the cash flow variance of a portfolio consisting of the sales revenue (or procurement cost) of an electric utility and the payoffs of derivatives on temperature and fuel price. In the first half of this section, the previous method [14] applied to the empirical analysis in this study will be briefly explained while supplementing the methodological and theoretical background. In the second half, we will elaborate on the spline basis functions used for non-parametric hedge models.

We first assume that electric utilities (hedgers) can use two types of temperature derivatives. One is the temperature "futures," whose payoff is defined as the observed temperature T_t minus its (predicted) seasonal trend $h_T(t)$ at date t (i.e., the prediction error $\varepsilon_{T,t} := T_t - h_T(t)$). As the temperature futures designed in this manner can be regarded as having an expected payoff value of 0 at the time of the prior derivative contract, they are practically easy to handle because, for example, they do not necessarily require premium payments between risk-neutral players.

The other is the temperature "derivatives" on the prediction errors, whose payoff is expressed as a form of an arbitrary function of $\varepsilon_{T,n}$ (if expressed as a univariate function, it will be a payoff function of the form $\psi(\varepsilon_{T,n})$). As with temperature "futures," temperature "derivatives" can also be designed so that the expected payoffs are 0; and this issue will be dealt with in Section 3.2. The main problem to be considered here is that electric utilities seek the optimal contract volumes of futures or payoff functions of derivatives, aiming to suppress their fluctuation risks of sales revenues.

3.1. Optimal Futures Contract Volume Calculation Problem

Of the two types of derivative products mentioned above, the hedging problem for temperature futures is considered in this subsection. For example, when an electric utility wants to hedge the daily fluctuation of sales revenue π_t with HH and temperature futures (whose payoffs are HH_t and $\varepsilon_{T,t}$, respectively), the minimum variance hedging problem to be solved is as follows:

$$\underset{f(\cdot)\in\mathcal{S}_{\lambda_f},\ \Delta(\cdot)\in\mathcal{S}_{\lambda_\Delta},\ \gamma(\cdot)\in\mathcal{S}_{\lambda_\gamma}}{\text{Min}} \text{Var}[\pi_t - f(t) - \Delta(t)HH_t - \gamma(t)\varepsilon_{T,t}] \tag{1}$$

where $\text{Var}[\cdot]$ denotes the sample variance; $f(t)$, $\Delta(t)$, and $\gamma(t)$ are the contract volumes of the discount bond, HH futures, and temperature futures, respectively, at date t; and $\mathcal{S}_\lambda$ is a set of smoothing spline functions, with the smoothing parameters λ that control the tradeoff between model fit and smoothness [15] (detailed in Section 3.3.1). Thus, (1) depicts a problem that minimizes the cash flow variance of a portfolio; it consists of the sales revenue, discount bonds, and futures on HH and temperature under the smoothing parameters. Importantly, this optimization problem corresponds to constructing the following prediction formula for π_t and applies a GAM to estimate the smoothing spline functions f, Δ, and γ:

$$\pi_t = f(t) + \Delta(t)HH_t + \gamma(t)\varepsilon_{T,t} + \eta_t \tag{2}$$

where η_t is the residual term with an average of 0. As proven in [10], estimating the GAM (2) corresponds to minimizing the variance of η_t under the smoothing conditions of f, Δ, and γ; hence, it is synonymous with solving (1) (here, η_t can be interpreted as the

"hedging error"). Note that the theoretical explanation for estimating the smoothing spline function using the GAM will be detailed in Section 3.3.

3.2. Optimal Derivative Payoff Calculation Problem

Considering that price has strong nonlinearity with respect to temperature, as explained in Section 2, this section introduces temperature "derivatives" with nonlinear payoff functions. Assuming the optimal payoff is given as a smooth function of temperature (when viewed on a specific date), changing smoothly according to the season, we formulate the following optimal payoff function calculation problem:

$$\underset{\widetilde{f}(\cdot)\in\mathcal{S}_{\lambda_{\widetilde{f}}},\ \Delta(\cdot)\in\mathcal{S}_{\lambda_{\Delta}},\ \widetilde{\psi}(\cdot)\in\mathcal{S}_{\lambda_{\psi 1},\ \lambda_{\psi 2}}}{\text{Min}} \text{Var}\left[\pi_t - \widetilde{f}(t) - \Delta(t)HH_t - \widetilde{\psi}(t,\ \varepsilon_{T,t})\right] \tag{3}$$

where $\widetilde{f}(t)$ is the contract volume of discount bonds and $\widetilde{\psi}(t,\ \varepsilon_{T,t})$ is the payoff function of weather derivatives, which does not include the date-dependent trend. Here, the payoff function $\widetilde{\psi}(t,\ \varepsilon_{T,t})$ is estimated as a two-dimensional tensor product spline function [23] (detailed in Section 3.3.2) with ANOVA decomposition [24] (see Appendix B) by applying the following GAM, as was done in (2):

$$\pi_t = \widetilde{f}(t) + \Delta(t)HH_t + \widetilde{\psi}(t,\ \varepsilon_{T,t}) + \eta_t \tag{4}$$

We can obtain the temperature derivative payoff, which does not include the deterministic date-dependent trend, by ANOVA decomposition, and the expected value can be regarded as 0 at each date t. In other words, the temperature derivatives here can be treated as those that may not require a premium payment at the contract time, as with the temperature futures introduced in the previous section.

3.3. Spline Function Estimation Procedure

In this section, to understand how the spline functions estimated by the GAM are defined and calculated, the bases of spline functions and their estimation algorithms are briefly explained.

3.3.1. Univariate Smoothing Spline Function

First, the univariate smoothing spline function is estimated as the function s that minimizes the penalized residual sum of squares (PRSS) given by:

$$\text{PRSS} = \sum_{n=1}^{N}\{y_n - s(x_n)\}^2 + J(s,\ \lambda),\ \text{where}\ J(s,\ \lambda) = \lambda\int\{s''(x)\}^2 dx. \tag{5}$$

In (5), the first term measures the approximation of the data, and the second term ("penalty term") $J(s,\ \lambda)$ adds penalties according to the curvature of the function. In this study, we estimate GAMs using the R 3.6.1 package "mgcv" to obtain the series of smoothing spline functions, wherein the smoothing parameter is calculated by general cross-validation criterion.

In particular, this study pays attention to the fact that different basis functions can be applied when estimating the function $s(x)$, wherein the basis represents the function $b_i(x)$ in the following formula:

$$s(x)\ = \sum_{i=1}^{k} b_i(x)\beta_i \tag{6}$$

where β_i is the coefficient of the basis function.

The basis functions (smoothing methods) have some variations, such as "cubic spline" (see, e.g., [19]), "cyclic cubic spline" [23], "P-spline" [21], and "thin plate spline" [20]. Of these, all bases other than "thin plate spline" are expressed as "piecewise polynomials" joined at the points called "knots." Each of them has the following characteristics:

- "Cubic spline": It is one of the most popular basis functions and is defined based on the third order "truncated power basis functions." In other words, it is expressed by cubic polynomials, and each piecewise polynomial smoothly connects at each knot (i.e., the value, the first derivative, and the second derivative are continuous; see Appendix C.1 for the concrete formulas).
- "Cyclic cubic spline": It has a basis function that is defined to be smoothly connected not only at each knot, but also at the start and end points of the domain (see Appendix C.2 for the concrete formulas). For this reason, it is suitable for robustly estimating the trend of periodic data.
- "P-spline": While using B-spline basis [25] (which is based on "a special parametrization of a cubic spline" [26]), P-spline uses the unique penalty term called "discrete penalty" [21]. Unlike the other basis functions having "continuous penalties" with "integral squared curvature," as shown in (5), P-spline penalizes the changes in discrete coefficients of adjacent bases of B-spline (see [21] for the formula of penalty term). Initially, B-spline bases are arranged so that adjacent bell-shape curves overlap each other (e.g., see "Figure 1" of [21]); therefore, even though the discrete penalty is imposed for coefficients of the bases, smoothness is ensured, as in the case of continuous penalties.
- "Thin plate spline": Also called "radial basis functions," the basis of the thin plate spline depends only on the distance (norm) from each control point rather than the coordinates for each dimension. Therefore, unlike the previous three bases, the "thin plate spline" does not have "knots" as connection points (see [20,23] for more detail).

In the R package "mgcv," "cubic spline" or "thin plate spline" is given as the default basis, and the implementer has the option of reselecting other bases. This study particularly examines the above-mentioned "cyclic cubic spline" and "P-spline," comparing them with these default bases; the result will be detailed in Section 5.2.

Note that although the "mgcv" package also implements dozens of smoothing methods, such as "adaptive smoothers," an extension of P-spline (see "smooth.terms" of [17]), the comparison in this study focuses on the above-mentioned (basic) four bases. Regarding those four bases, Wood [17], who implemented "mgcv", introduces them as representatives of smoothers and provides detailed explanations (see the first part of "Section 5.3" in [17]), and there are many applied researches compared to other bases; therefore, we decided to choose those basis functions. That is, this study aims to explore the degree of improvement caused by reselection of the bases, and identifying the best basis functions is a future task.

Note also that there are multiple applied research cases in fields such as meteorology for "cyclic cubic spline," which is suitable for modeling periodic trends, but as far as we know, previous research applied in the field of energy does not exist, excluding the literature [27] that modeled electricity demand. Hence, this study is the first attempt to apply the "cyclic cubic spline" to model periodic electricity prices (for which trigonometric-function-based Fourier series expansion has been adopted in many studies).

3.3.2. Multivariate Smoothing Spline Function

Next, we describe the multivariate smoothing spline function. When spline functions are extended to multiple dimensions, their smoothing approaches are roughly divided into "tensor product smoothing" (tensor product spline) and "isotropic smoothing" [23].

3.3.2.1. Tensor Product Smoothing (Tensor Product Spline)

The tensor product spline function has different basis functions for each dimension, and the basis is given by the tensor product. For example, in the case of the bivariate function $s(x, z)$, the tensor product spline is written as the sum of the products of the basis functions, such as $a_i(x)$ and $b_j(z)$, as follows (note that for $a_i(x)$ and $b_j(z)$, it is possible to specify the different types of basis separately [17,23]):

$$s(x, z) = \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \beta_{i,\,j} a_i(x) b_j(z). \tag{7}$$

For estimating the function $s(x,z)$, the following penalty term $J_{te}(s,\ \lambda_x,\ \lambda_z)$ is included in the PRSS (i.e., (5) for the univariate case) that should be minimized [23]:

$$J_{te}(s,\ \lambda_x,\ \lambda_z) = \int_{x,z} \lambda_x \left(\frac{\partial^2 s}{\partial x^2}\right)^2 + \lambda_z \left(\frac{\partial^2 s}{\partial z^2}\right)^2 dxdz. \tag{8}$$

3.3.2.2. Isotropic Smoothing

Among the multivariate spline functions, the concept opposite to the tensor product spline is "isotropic smoothing" [23], and its representative one is the thin plate spline (note that there are other isotropic smoothing approaches, such as "Duchon splines" [28], which are a generalization of thin plate splines, and "soap film smoothing" [29], which is based on the idea of constructing a 2-D as smooth as a film of soap). As mentioned in the previous section, the basis of the thin plate spline is given as a function based on the distance (norm) from specific points, so when it is extended in multiple dimensions, the penalty term is different from that of the tensor product spline. For example, the penalty term $J_{tp}(s,\ \lambda)$ for the bivariate thin plate spline function is given as the following Equation [23]:

$$J_{tp}(s,\ \lambda) = \lambda \int_{x,z} \left(\frac{\partial^2 s}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 s}{\partial x \partial z}\right)^2 + \left(\frac{\partial^2 s}{\partial z^2}\right)^2 dxdz. \tag{9}$$

As is clear from the comparison between (8) and (9), the term related to the mixed partial derivative $\partial^2 s/\partial x \partial z$ is added to the penalties for the thin plate spline. The thin plate spline is characterized by the isotropic addition of smoothing penalties at each point, as referenced by its name, which comes from its resemblance to the bent shape of a thin elastic plate.

While the thin plate spline is suitable if different axis units are the same, the tensor product spline, which can independently incorporate the smoothing conditions for each axis, is more suitable if the axis units are different. Therefore, this study adopts the tensor product spline for the pricing of the derivative $\widetilde{\psi}(t,\ \varepsilon_{T,t})$, with smooth trends in directions with different units such as date and temperature.

4. Construction of Hedging Models

In this section, we construct the concrete hedge models using the methods introduced in Section 3. Because the models have the same forms as the ones used in [14], only the outline is provided in this section.

4.1. Base Model Consisting of Fuel Price and Calendar Trend

First, considering that PJM electricity prices are strongly linked to HH prices and day type, and that they have annual change trends, as explained in Section 2, the following GAM is constructed, referred to as the "base model":

$$\text{where } \begin{cases} \pi_t = f(t) + \Delta(t) HH_t + \eta_t \\ f(t) := f_O(t) + f_H(t) I_{H,t} + f_P(t) Period_t \\ \Delta(t) := \Delta_O(t) + \Delta_P(t) Period_t \end{cases} \tag{10}$$

where f and Δ are yearly cyclical trends estimated as spline functions by the GAM (as $f(Seasonal_t)$ with yearly cyclical dummy variables $Seasonal_t\ (= 1, \ldots, 365 (or\ 366))$ [30], denoted as $f(t)$ and $\Delta(t)$ for concise notation), $I_{H,\,t}$ is a dummy variable for holiday, and $Period_t$ is the elapsed day of date t (annualized) from the beginning of the starting year of the data. Of these, the term $f_P(t)Period_t$ ($\Delta_P(t)Period_t$) is introduced because the calendar trend (sensitivity of HH to π_t) is assumed to have yearly cyclical trends, even when viewed at the rate of annual change. When estimating the model, three sets of the same data sample are used side by side so that the start and end points of the estimated cyclical trends f and Δ are approximately connected; by doing so, the desired yearly cyclical trends can be obtained as the estimated function in the middle domain (see "Appendix" in [14]).

4.2. Temperature Futures

Next, for the case wherein temperature futures can be used, the following hedging model is considered:

$$\pi_t = f(t) + \Delta(t)HH_t + \gamma_1(t)\varepsilon_{Tmin,t} + \gamma_2(t)\varepsilon_{Tmax,t} + \eta_t \tag{11}$$

where $\varepsilon_{Tmin,t}$ and $\varepsilon_{Tmax,t}$ are the payoffs of minimum and maximum temperature futures on day t, respectively, and $\gamma_\cdot(t)$ is the estimated yearly cyclical trend, corresponding to the contract volume of those futures.

4.3. Temperature Derivatives Estimated by the Tensor Product Spline

Similarly, a hedging model for temperature derivatives is constructed as follows:

$$\pi_t = f(t) + \Delta(t)HH_t + \widetilde{\psi_1}(t, \varepsilon_{Tmin,t}) + \widetilde{\psi_2}(t, \varepsilon_{Tmax,t}) + \eta_t \tag{12}$$

where $\widetilde{\psi}_\cdot(t, \varepsilon_{\cdot,t})$ are the smooth payoff functions of temperature derivatives, which change smoothly depending on the date t, estimated as tensor-product spline functions from which yearly cyclical trends have been removed via ANOVA decomposition, as explained in Section 3.2. Notably, the two payoff functions on minimum and maximum temperature derivatives can be uniquely estimated here because the date-dependent trend is unified into the identical term $f(t)$ through ANOVA decomposition. (More specifically, if ANOVA decomposition is not applied, f, ψ_1, ψ_2 all contain trends with respect to t (i.e., they have overlapping degrees of freedom with respect to t), and each function shape cannot be determined. Conversely, in (12) with ANOVA decomposition applied, the trend for t is removed from the derivative payoff functions $\widetilde{\psi_1}$, $\widetilde{\psi_2}$ and is explained only by f, which solves the problem of overlapping degrees of freedom. See also Appendix B for details on ANOVA decomposition.)

4.4. Temperature Derivatives for the Squared Prediction Error

The temperature derivatives' payoffs in (12) are estimated differently by each hedger, but here, we consider a "standard" derivative on temperature, which can be commonly traded by multiple hedgers. From the idea of approximating $\widetilde{\psi}(t, \varepsilon_{\cdot,t})$ with a quadratic function for $\varepsilon_{\cdot,t}$, we introduce temperature derivatives on squared prediction errors and construct the following model:

$$\pi_t = f(t) + \Delta(t)HH_t + \gamma(t)\varepsilon_{T,t} + \tau_1(t)\left({\varepsilon_{Tmin,t}}^2 - \overline{{\varepsilon_{Tmin,t}}^2}\right) + \tau_2(t)\left({\varepsilon_{Tmax,t}}^2 - \overline{{\varepsilon_{Tmax,t}}^2}\right) + \eta_t. \tag{13}$$

where ${\varepsilon_{\cdot,t}}^2 - \overline{{\varepsilon_{\cdot,t}}^2}$ are the payoffs of the squared prediction error derivatives on temperatures ($\overline{{\varepsilon_{\cdot,t}}^2}$ is the predicted value (sample mean) of ${\varepsilon_{\cdot,t}}^2$), and $\tau_\cdot(t)$ are the spline functions representing the contract volumes of the derivatives estimated by the GAM. These temperature squared error derivatives also have zero expected payoffs for each t, and similar to other derivatives, they do not require premium payments.

5. Empirical Analysis

This section empirically validates the hedging models introduced in Section 4 by using PJM market data. First, Section 5.1 applies those models to each of the three different "business risk models" and compares hedge effects as well as the shapes of the estimated derivative payoffs. Then, Section 5.2 compares the hedge effect by using the different bases described in Section 3.3.1. The empirical data used are as follows:

(a) Demand $D_{t,\,h}$ (TWh): the hourly load of the entire PJM-RTO [31].
(b) Electricity price $S_{t,\,h}$ (USD/MWh): day-ahead hourly spot price of PJM-RTO [31].
(c) Maximum and minimum temperatures $Tmax_t$, $Tmin_t$: population-weighted average of four main cities (Philadelphia, Pittsburgh, Baltimore, and Newark) published by the National Oceanic and Atmospheric Administration (NOAA) [32].

(d) Henry Hub natural gas price HH_t (USD/million BTU): historical daily HH spot price FOB [33].

The model parameters and functions are estimated from the in-sample period data (1 January 2011–31 December 2017), and the hedge effects are calculated from the out-of-sample data (1 January 2018–31 December 2018). In this study, we choose only the above four major cities in the PJM area (as with [34] for example) for the temperature index, but it may be possible to construct a more fitted hedge model (i.e., obtain higher hedge effect) by increasing the number of temperature observation points. However, it should be also noted that when assuming that the temperature derivatives are traded in practice, the smaller the number of points, the easier it is for traders to understand and handle.

5.1. Empirical Analysis by Business Risk Models

This section verifies estimated trend functions and hedge effects in the context of comparing business risk models. As introduced in Section 1, we deal with the three business risk models as exposed to (i) both price and demand risks, (ii) price risk only, and (iii) demand risk only. That is, for each case, the hedged target (i.e., the hedger's fluctuating revenue/cost) is expressed, respectively, as (i) the sum of the product of hourly spot price $S_{t,\,h}$ and demand $D_{t,\,h}$($\pi_t = \sum_h S_{t,\,h} \times D_{t,\,h}$, referred to as the "product model"); (ii) price ($\pi_t = (1/24)\sum_h S_{t,\,h}$; "price model"); and (iii) demand ($\pi_t = \sum_h D_{t,\,h}$; "demand model").

Note that because the superiority of the cyclic cubic spline over different bases is revealed by the empirical analysis in Section 5.2, this section uses that basis to compare business risk models.

5.1.1. Trend Estimation of Hedge Models

5.1.1.1. Optimal Payoff Function of the Temperature Derivatives

First, Figure 3 displays the min/max temperature derivatives' payoff functions $\widetilde{\psi_1}(t, \varepsilon_{Tmin,t})$ and $\widetilde{\psi_2}(t, \varepsilon_{Tmax,t})$ of the "product model," which were simultaneously estimated as a tensor product spline function using ANOVA decomposition in hedge model (12). In both cases, it can be confirmed that the trends in the seasonal direction are removed (e.g., having shapes with zero mean at each date t) via ANOVA decomposition. In addition, the payoff of the derivative of minimum temperature (corresponding to the sensitivity of the minimum temperature prediction error to the sales revenue) is specifically increased as temperatures drop in winter, whereas that of maximum temperature is increased as temperatures rise significantly in summer, reflecting that both distinctive effects complement each other.

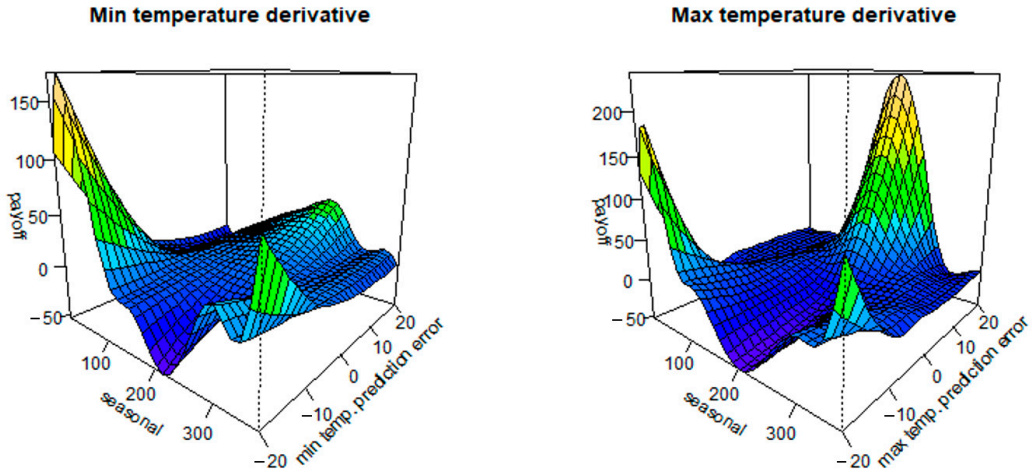

Figure 3. Estimated temperature derivatives' payoffs in the product model (12).

Next, Figure 4 shows the payoff functions for the temperature derivatives of the "price model." The shapes of the payoff functions are not significantly different from those of

the product model shown in Figure 3, so it can be inferred that the nonlinearities of the "product" model's derivatives mostly result from those of the "price" to the temperature (to put it in detail, the daily change in the slope of the maximum temperature derivative during summer in Figure 3 is slightly more rapid than that in Figure 4, which may indicate that the product model has slightly stronger nonlinearity than the price model).

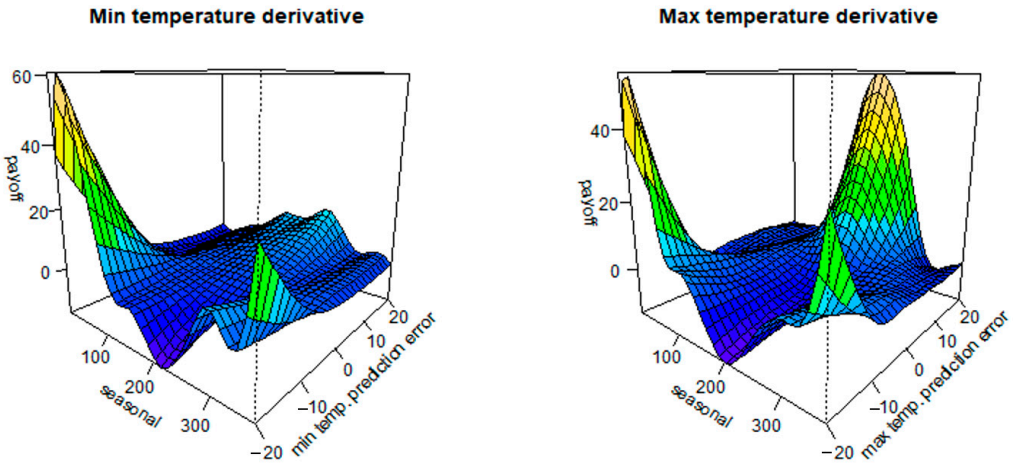

Figure 4. Estimated temperature derivatives' payoffs in the price model (12).

The payoff functions for the temperature derivatives of the demand model are shown in Figure 5. Similar to that of the product model, the slope of the payoff function is positive in summer and negative in winter, but notably, the payoff function of the demand model is smoother than that of the product model (i.e., the nonlinearity is relatively small). This probably indicates that the temperature sensitivity to the demand with the normal temperature on a specific date is relatively small, as seen in Figure 1.

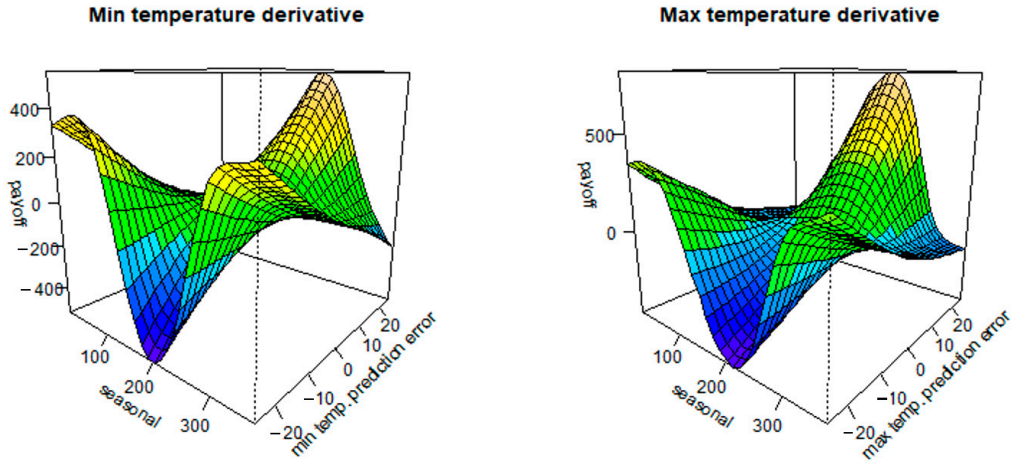

Figure 5. Estimated temperature derivatives' payoffs in the demand model (12).

5.1.1.2. Optimal Contract Volume of the Squared Temperature Prediction Error Derivatives

Figure 6 displays the estimated contract volume trends (the dotted line indicates 95% confidence interval) of the squared prediction error derivatives of the min/max temperatures in the "product model" (13). They can be rephrased as trends that reflect the magnitudes of the (downward) convexities of temperature sensitivities to the sales revenue. Both trends rise in summer and winter, indicating that the maximum temperature's nonlinearity to sales revenue is particularly strong in summer, while that of the minimum temperature is particularly strong in winter (consistent with the non-parametrically-priced derivatives shown in Figure 3). Similarly, Figure 7 shows the same trends for the price

model. The shapes are not significantly different from the product model shown in Figure 6, probably because of the same reason explained in Section 5.1.1.1.

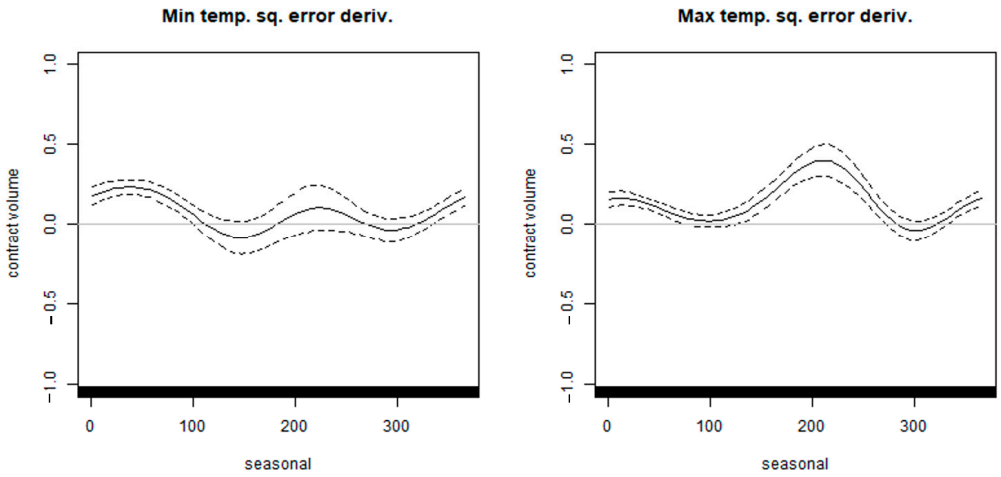

Figure 6. Estimated contract volume of the temperature squared error derivatives in the product model (13).

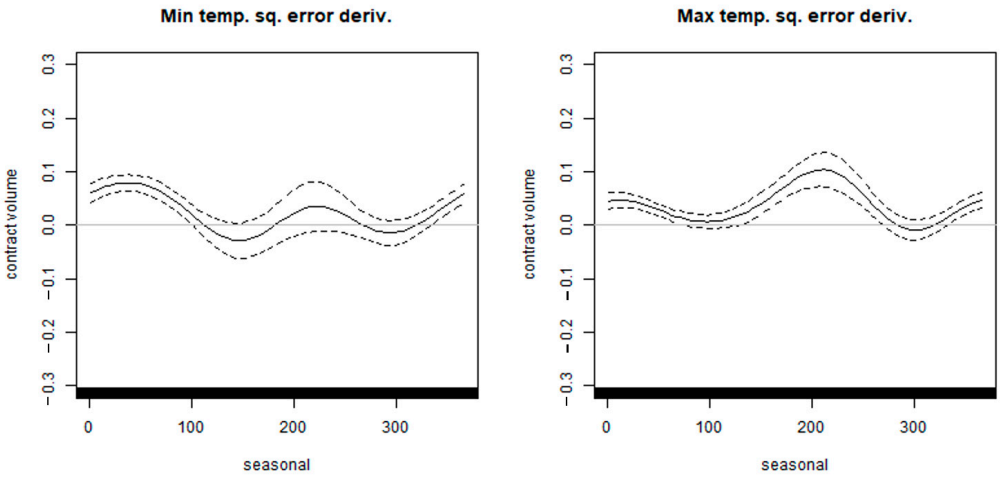

Figure 7. Estimated contract volume of the temperature squared error derivatives in the price model (13).

Figure 8 shows the estimated contract volume of the temperature squared error derivatives in the demand model. The derivative contract volumes of the minimum and maximum temperatures approaching 0 is common in winter, but in summer, that of the maximum temperature is significantly higher while that of the minimum temperature approaches 0. This may be due to the following reasons: In the PJM area, the absolute value of the temperature sensitivity (slope) to power demand tends to be higher in summer than in winter, as seen in Figure 1, and the "change rates" in temperature sensitivity have a similar seasonal tendency (i.e., convexity to temperature). Probably reflecting these tendencies, the maximum temperature strongly explains such convexity around the summer season, while the minimum temperature complementarity explains the (rest of the) seasonal changes in other convexities. In addition, this estimated result is consistent with the fact that the downward convexity of the maximum temperature derivative payoff becomes relatively larger during the summer, as confirmed in Figure 5.

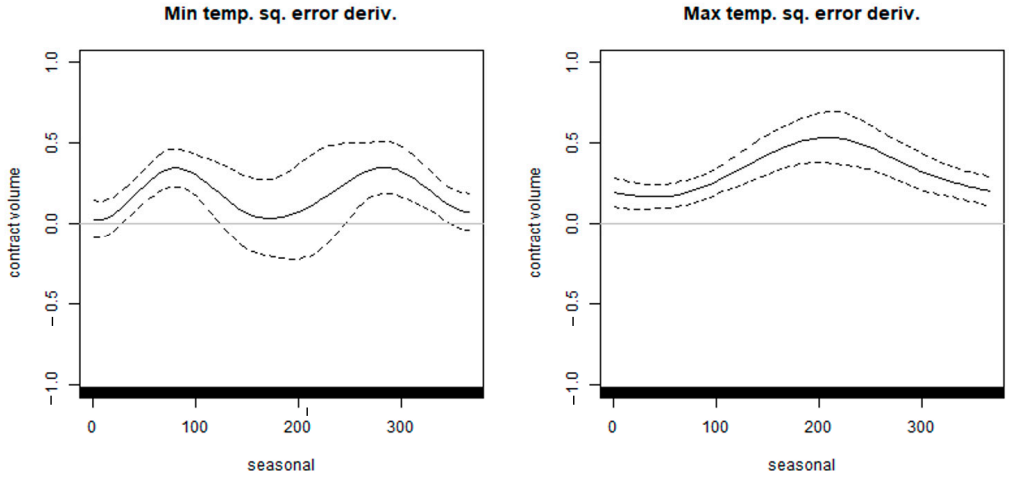

Figure 8. Estimated contract volume of the temperature squared error derivatives in the demand model (13).

5.1.2. Measurement of Hedge Effects

In the following, we measure the hedge effect of each derivative for each business risk model. This study uses the variance reduction rate (VRR), defined as follows, and it refers to 1-VRR as the hedge effect:

$$\text{VRR} := \frac{\text{Var}[\text{hedge error of the target model}]}{\text{Var}[\text{hedge error of the base model}]}. \tag{14}$$

5.1.2.1. Cumulative/Individual Hedge Effects by Derivatives

Here, we analyze changes in hedge effects when each derivative (hedge model term) is cumulatively combined; the results are summarized in Figures 9–11 for "product model," "price model," and "demand model," respectively. Each figure illustrates the "single contribution ratio" when each derivative is used alone (bar graph), the "cumulative contribution ratio" (corresponding to the R-squared; see, e.g., ref. [35] for "out-of-sample R-squared statistic") when the derivatives are combined in order from the top (blue line graph), and the "cumulative hedge effect" of the temperature derivatives compared with the "base model" (3) (red line graph). All are measured for the three models (product, price, and demand models).

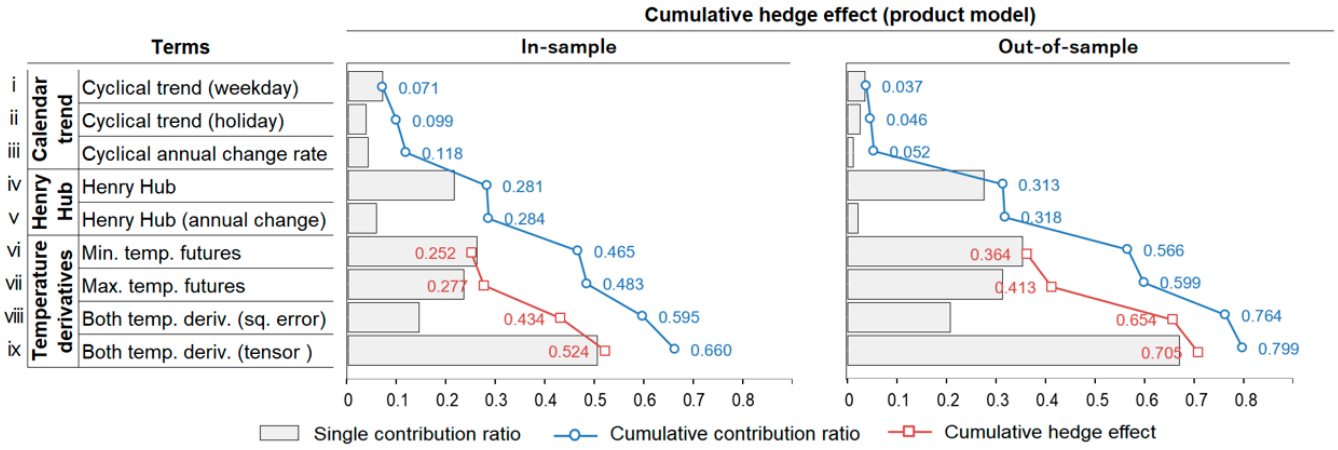

Figure 9. Contribution ratio and hedge effect of product model.

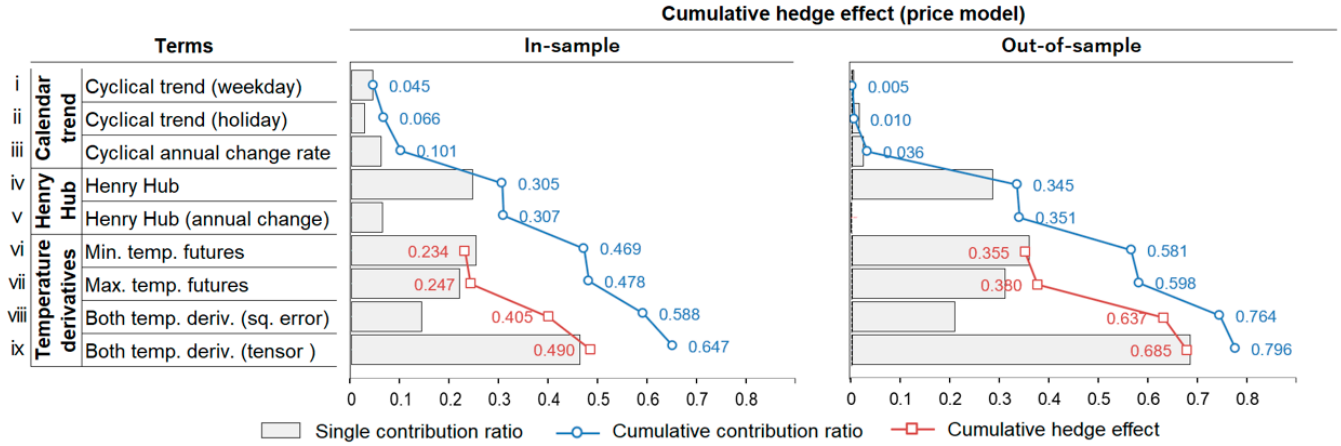

Figure 10. Contribution ratio and hedge effect of price model.

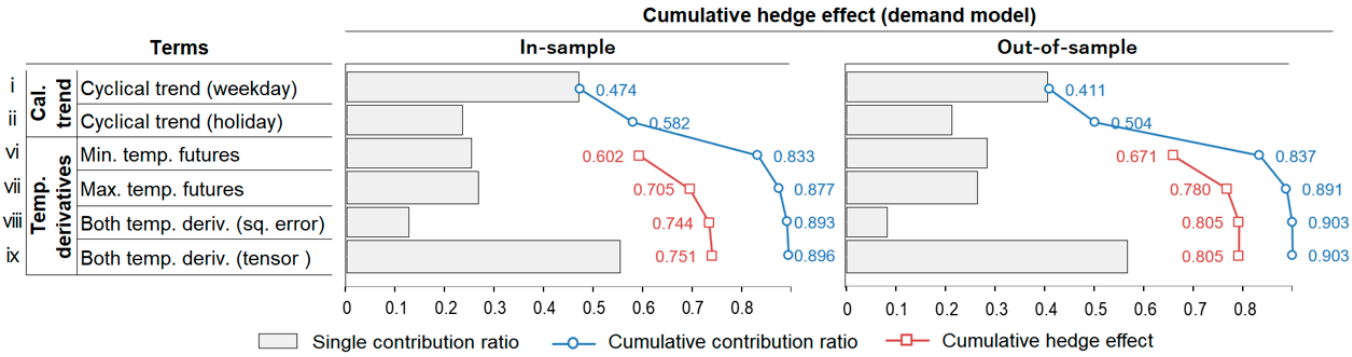

Figure 11. Contribution ratio and hedge effect of demand model.

Note that the Roman numerals of the terms in Figures 9–11 correspond to the following: i. $f_O(t)$; ii. $f_H(t)I_{H,t}$; iii. $f_P(t)Period_t$; iv. $\Delta_O(t)HH_t$; v. $\Delta_P(t)Period_tHH_t$; vi. $\gamma_1(t)\varepsilon_{Tmin,t}$; vii. $\gamma_2(t)\varepsilon_{Tmax,t}$; viii. $\tau_1(t)({\varepsilon_{Tmin,t}}^2 - \overline{\varepsilon_{Tmin,t}^2}) + \tau_2(t)({\varepsilon_{Tmax,t}}^2 - \overline{\varepsilon_{Tmax,t}^2})$; ix. $\widetilde{\psi}_1(t, \varepsilon_{Tmin,t}) + \widetilde{\psi}_2(t, \varepsilon_{Tmax,t})$ (in the cumulative usage case including up to ix, terms vi–viii are excluded). Note also that in the demand model, terms regarding HH are excluded since no correlation is assumed.

First, for all three models, the temperature derivatives had the highest single contribution ratios (around 60–70%) among all derivatives (terms) in the out-of-sample period. Next, the cumulative contribution ratios increased monotonously with the inclusion of each term, even in the out-of-sample case, and they reached close to 80% for both the product and price models and over 90% for the demand model. Similarly, the cumulative hedge effects increased monotonically and reached approximately 70% for both the product and price models, and over 80% for the demand model. When the maximum temperature futures were combined with the minimum temperature futures, the hedge effect was improved by approximately 3–5 percentage points for the product and price models and approximately 11 percentage points for the demand model, respectively. Hence, it is suggested that the two different temperature products have complementary effects for all models.

Regarding the product and price models, when compared with using only temperature futures, the combined use of the squared error temperature derivatives further improved the hedge effect by about 24–26 percentage points, and further improvement of approximately 5 percentage points occurred when using the derivatives of the tensor product splines. This result reflects the strong nonlinear correlation between temperature and PJM price. (This may be easier to understand when the price (product) is simply regressed by temperature regardless of date t. In such case, the quadratic function (corresponding to the payoff of squared error temperature derivatives) fits better than the linear function (corresponding to the payoff of temperature futures), and arbitrary spline function

(corresponding to the payoff of tensor product spline derivatives) fits further better than both, thereby reducing the variance of the residuals (i.e., variance of hedging error; see the explanation of (2)). At this time, the numerator of VRR (4) becomes smaller, so the hedging effect (1-VRR) becomes larger.

On the other hand, regarding the demand model, the high hedge effect of 67% was confirmed only by using the minimum temperature futures with a linear payoff function. Presumably, it indicates that the temperature sensitivity to demand has a small nonlinearity, as was confirmed with the derivative payoff shape in Figure 5. Regarding the hedge effect of nonlinear derivatives, a 2 percentage-point improvement was confirmed for the squared error derivatives, but no improvement was seen for the tensor-product derivatives. This result implies that "customized yet standardized" square error derivatives (combined with temperature futures) may be superior to "made-to-order" tensor-product derivatives, in that the standardized derivatives allow for liquid trading.

In this way, the squared error derivative significantly improves the hedging model of both the product and price models, and improves even the hedge effect of the demand model, which has relatively weak nonlinearity. It is suggested that for many risk types, electric businesses may be able to trade it in common for efficiently hedging "nonlinearity-derived" fluctuation risks. Note that although this study verified the hedging effects of different types of "electric utilities," since the payoff function of "customized yet standardized derivative" is defined only by the (public) measured temperature, it could also be used for businesses in other sectors affected by weather, such as the agriculture and leisure industries.

5.1.2.2. Monthly Hedge Effect

Figure 12 demonstrates the monthly hedge effects (1-VRR) of the temperature derivatives (estimated by the tensor-product spline functions) for each of the three models for both in-sample and out-of-sample periods. For each period, the hedge effects have generally similar seasonal tendencies. The hedge effect tends to increase in summer (June–July) and winter (December–January) for each model, which corresponds to the payoff functions of each model's derivatives having an extremely steep slope during these seasons, as seen in Figures 3–5. The demand model has a higher hedge effect throughout the period (for all months) than the price or product model. It is suggested that the electricity demand tends to fluctuate relatively greatly because of temperature. The hedge effect for the price and product models is not so large during spring and autumn (around April–May or September–November) because the effect of temperature on price fluctuations during these periods is smaller than other factors, such as changes in the market environment and power supply operation (e.g., in the price model of September 2018, the hedge effect of the out-of-sample period has a negative value of -0.55). Hence, derivative trading strategies limited to summer and winter may be effective for price risk and product risk.

See Appendix C as well, wherein we verify the accuracy of the model by month from the perspective of "hedge error."

5.2. Comparison between Basis Functions

This section examines the extent to which the hedge effects measured in Section 5.1.2 may change when using different basis functions, as introduced in Section 3.3.1. The examined basis functions include: (a) thin plate spline / cubic spine ("tp/cr"), which is the default case of the R package "mgcv" (wherein the thin plate spline is used for univariate splines, and the cubic spline is used for tensor product splines); (b) P-spline ("ps"); and (c) cyclic cubic spline ("cc"). We target all hedge cases involving the cumulative hedge effect of the weather derivative seen in Figures 9–11, and we compare the basis functions by changing only those for seasonal trends in the date direction in common for all cases (i.e., for the temperature direction in the tensor product spline, the default basis "cr" is used as is in all cases). For each of these bases, we compare the cumulative hedge effect

and contribution ratio for all business risk models. The result is shown in Table 1 (the red gradation is colored by comparing the three values of each basis within the same model).

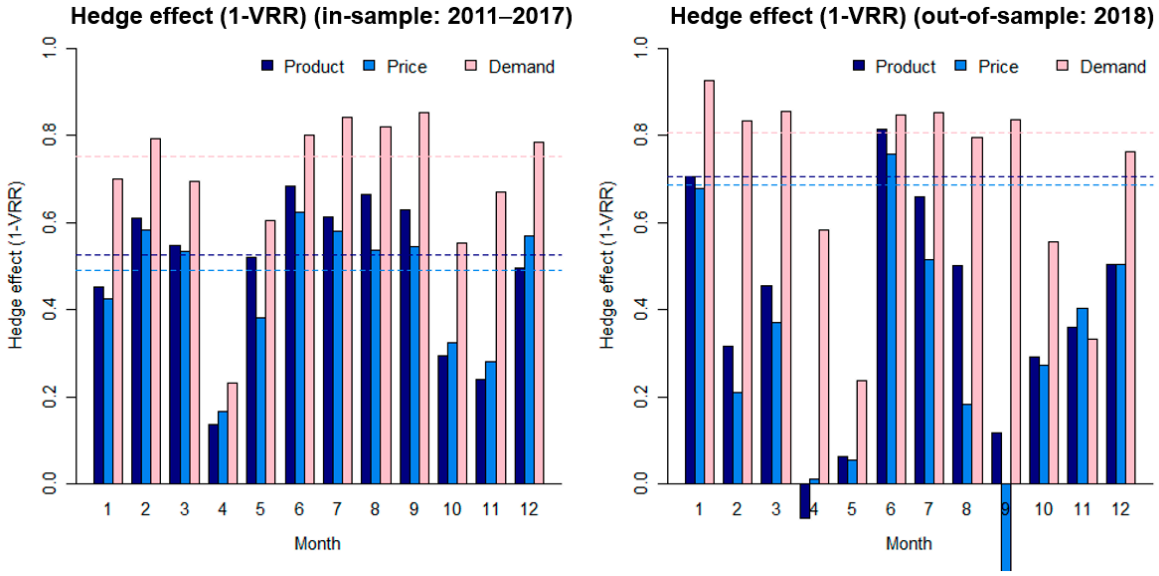

Figure 12. Monthly hedge effect for each business risk model.

Table 1. Comparison of cumulative hedge effect/cumulative contribution ratio by basis.

Business Risk Model		In Sample						Out-of-Sample							
		Cumulative Contribution Ratio			Cumulative Hedge Effect			Cumulative Contribution Ratio			Cumulative Hedge Effect			Δ (to tp/cr)	
		tp/cr	ps	cc	tp/cr	ps	cc	tp/cr	ps	cc	tp/cr	ps	cc	ps	cc
Product	vi	0.480	0.477	0.465	0.267	0.262	0.252	0.521	0.546	0.566	0.238	0.289	0.364	21%	53%
	vii	0.495	0.492	0.483	0.289	0.283	0.277	0.548	0.571	0.599	0.281	0.329	0.413	17%	47%
	viii	0.601	0.598	0.595	0.438	0.433	0.434	0.720	0.722	0.764	0.555	0.565	0.654	2%	18%
	ix	0.660	0.641	0.660	0.520	0.493	0.524	0.762	0.788	0.799	0.621	0.668	0.705	8%	13%
Price	vi	0.484	0.483	0.469	0.246	0.244	0.234	0.545	0.568	0.581	0.248	0.290	0.355	17%	43%
	vii	0.494	0.491	0.478	0.259	0.257	0.247	0.564	0.585	0.598	0.278	0.318	0.380	14%	37%
	viii	0.596	0.592	0.588	0.409	0.404	0.405	0.731	0.733	0.764	0.555	0.561	0.637	1%	15%
	ix	0.647	0.633	0.647	0.484	0.463	0.490	0.781	0.782	0.796	0.637	0.642	0.685	1%	8%
Demand	vi	0.837	0.826	0.833	0.607	0.589	0.602	0.835	0.837	0.837	0.668	0.667	0.671	0%	0%
	vii	0.881	0.870	0.877	0.712	0.694	0.705	0.887	0.894	0.891	0.772	0.783	0.780	1%	1%
	viii	0.896	0.888	0.893	0.748	0.735	0.744	0.902	0.907	0.903	0.802	0.810	0.805	1%	0%
	ix	0.898	0.891	0.896	0.753	0.744	0.751	0.903	0.905	0.903	0.804	0.804	0.805	0%	0%

As can be seen from this table, the value of "tp/cr" in all cases is the highest of the three basis functions for the in-sample period, but it is the lowest for the out-of-sample period overall. On the contrary, "cc" tends to have a small value for the in-sample, but a high value in the out-of-sample period. Moreover, "ps" is almost in the middle of the two. In other words, such a phenomenon of value reversal between the in-sample and the out-of-sample suggests that "cc" and "ps" are superior to "tp/cr" in terms of robustness. Looking at the improvement of the hedge effect of "cc" and "ps" in particular, when compared with the default case "tp/cr" in the out-of-sample (shown in the right two columns), relatively large improvements are observed for the "product model" (and the "price model") with strong nonlinearities (note that in Section 5.1.2.1, when measuring the improvement of the cumulative hedge effect by additionally incorporating hedge products, it was meaningful to use "percentage point," which measures the differentials of the hedge effects; however, in this section, the improvement "ratio" from the default case is measured to compare the

methodologies in the same hedge model). In particular, "cc" has larger hedge effects than the default basis in all cases, and in the product model, it improves by no less than 13–53% (Note 1: The results of the product model in Table 1 can be easily verified by using the R source code published in [14] and by only changing each of the default basis functions to "ps" or "cc"; Note 2: The values of the hedge effect shown in [14] exactly match the values in the crossing sells of the "tp/cr" columns and the "product" rows in Table 1; Note 3: Although not shown in this table, when the "normal" cubic spline "cr" was used instead of the thin plate spline "tp," the change was very slight in all cases, and in fact, more cases worsened).

The technical consideration of the above results is as follows. First, the high robustness of the P-spline is likely because the "discrete penalty" of the P-spline contributes to the avoidance of overfitting. In fact, the P-spline has an advantage in that the "loss of control" problem that tends to occur when using "continuous penalty" can be avoided [36], and such a mechanism may have worked. Regarding the cyclic cubic spline, its high robustness probably comes from the constraint that the start and end points are smoothly connected. The hedge model method proposed in [14] tried to impose this constraint through a data-driven manner using three sets of the same data sample side by side, so that the start and end points of the yearly cyclical trend were smoothly connected (see, e.g., "Appendix" in [14]). However, we found in this study that such ingenuity is not always sufficient, and that by incorporating similar constraints as well into the basis functions used in the model, robustness is further ensured, and the out-of-sample hedge effect can be significantly improved. More interestingly, in the comparison among business risk models, the improvement resulting from appropriate basis selection is remarkably large, especially for the product model (and price model) with strong nonlinearity, and small for the demand model with weak nonlinearity. This result suggests that the stronger the nonlinearity inherent in the model, the more important the robustness (constraint strength) to be incorporated into the functional expression (Appendix D).

Although the basis function selection may be overlooked by modeling practitioners, it may be a highly critical issue in robustly estimating a model in which strong nonlinearities are intertwined in a complex manner, such as the hedging models treated in this study.

6. Conclusions

In this study, paying attention to the fact that different types of electric utilities are exposed to risks of demand, price, and both, we verified the hedge effects for each of the three business risk models ("demand model," "price model," and "product model") using a previously proposed temperature derivative portfolio estimated using non-parametric hedging models. In addition, we found that choosing the appropriate basis for spline function can ensure the robustness of the model and significantly improve the out-of-sample hedge effects.

First, regarding the comparison between the three business risk models, the following empirical results and suggestions were obtained:

- The nonlinearity of the temperature derivative payoffs by the business risk model is strong in the product model and the price model, and relatively weak in the demand model.
- Reflecting this, the non-parametrically-priced derivative payoff function, which can flexibly express strong nonlinearity, has the highest hedge effect on the product model and the price model.
- On the contrary, for the demand model, the hedge effect of the non-parametric derivatives does not exceed that of the standard derivative on the squared temperature prediction error; hence, the squared error derivatives may be superior in that it allows for liquid trading.

It was confirmed that the squared error derivative has high hedge effects on both product and price models, which are comparable to the non-parametric derivatives. This result also suggests that this "customized yet standardized" squared error derivative

is promising as a new standard derivative that can be traded between players exposed to many different business risks. On the other hand, it is also true that the "made-to-order" non-parametric derivatives have the highest hedge effect for both product and price models. Therefore, in practical decision-making scenes, derivative contracts may need to be made after considering the trade-off between liquid tradability and maximization of the hedge effect.

Next, regarding the basis selection for the spline functions, we obtained the following implications:

- When estimating the series of trends existing in the hedge models, using the P-spline or the cyclic cubic spline instead of the thin plate spline or cubic spline set as the default in the R "mgcv" package can secure the robustness of the models; as a result, the out-of-sample hedge effect may be significantly improved.
- The improvement of the hedge effect by appropriate basis selection is larger for the product and price models than for the demand model. This means that the stronger the nonlinearity in the model, the more critical the basis selection that can robustly express the inherent trends of the data.

When the cyclic cubic spline was applied to the seasonal trend of the "product model" with the strongest nonlinearity, surprisingly, the hedge effect improved by 13–53%, compared with the previous empirical results using the default case that was demonstrated in [14]. Although the selection of basis functions seems to be overlooked in practice, we conclude that it is extremely important to keep in mind for the robust estimation of models with strong nonlinearities, as treated in this study.

The non-parametric hedging models we have proposed have been evolving in demonstrating applicability to different empirical data and devising methodologies for ensuring robustness. In the electricity market of the future, wherein transactions of decentralized players are assumed to increase significantly, it is expected that there will be increasing needs for financial instruments that can flexibly hedge fluctuation risks in finer time granularity (daily and hourly), such as the weather derivatives used in this study. Our future task is to expand the empirical analysis further and refine the model aiming for the practical application of this unique non-parametric hedging model and the high-resolution weather derivatives.

Author Contributions: Conceptualization, Y.Y.; methodology, T.M.; software, T.M.; validation, T.M.; data curation, T.M.; writing—original draft preparation, T.M.; writing—review and editing, Y.Y.; visualization, T.M.; supervision, Y.Y.; project administration, Y.Y.; funding acquisition, T.M. and Y.Y. Both authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by a Grant-in-Aid for Scientific Research (A) 16H01833, Grant-in-Aid for Scientific Research (A) 20H00285, Grant-in-Aid for Challenging Research (Exploratory) 19K22024, and Grant-in-Aid for Young Scientists 21K14374 from the Japan Society for the Promotion of Science (JSPS).

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

Nomenclature

t	Date
h	Hour
π_t	Sales revenue (or procurement cost) of an electric utility
$D_{t,\,h}$	Demand
$S_{t,\,h}$	Spot price
HH_t	Henry Hub natural gas price
T_t	Temperature
$Tmax_t$, $Tmin_t$	Maximum/Minimum temperature

$\varepsilon_{T,t}$	Payoff of the temperature futures (temperature prediction error)
${\varepsilon_{T,t}}^2 - \overline{{\varepsilon_{T,t}}^2}$	Payoff of the squared prediction error derivative on temperature
$\psi(t, \varepsilon_{T,t})$, $\widetilde{\psi}(t, \varepsilon_{T,t})$	Payoff of the temperature derivative estimated by tensor product spline function
$f(t)$, $\widetilde{f}(t)$	Contract volume of discount bonds
$\Delta(t)$	Contract volume of HH futures
$\gamma(t)$	Contract volume of the temperature futures
$\tau(t)$	Contract volume of the squared prediction error derivatives on temperature
η_t	Residual term (hedging error of derivative portfolio]
λ	Smoothing parameter
$\mathcal{S}_\lambda$	A set of smoothing spline functions with smoothing parameter λ
$s.(x)$	Spline function
J	Penalty term of penalized residual sum of squares
$b_i(x)$	Basis functions of spline function
β_i	Coefficients of the basis functions

Appendix A. Transaction Flow of "Customized Yet Standardized" Derivatives

Figure A1 shows the transaction flowchart of the "customized yet standardized" derivatives. Although the figure is created for an electric utility (retailer) exposed to "product risk," if it is not exposed to price or demand risk, the flow excluding (fixing) that risk may be considered. The transaction procedure is as follows:

1. The electric utility optimizes the contract volume (such as $\gamma(t)$ and $\tau(t)$) of the standard derivatives for each future delivery date based on the past profit/loss function (for the retailer in the figure, it is procurement cost $\pi_{t,\,h} := S_{t,\,h} \times D_{t,\,h}$).
2. The utility makes a contract (transaction) of the temperature derivatives of the volume calculated in 1. in the derivative market. At this time, no premium payment is made.
3. The utility purchases electricity at the spot market and pays the corresponding procurement cost (or records the cost in its account books) on the day before the delivery date of electricity.
4. The utility receives (or pays) a payoff of temperature derivatives calculated based on the measured temperature on the delivery day of electricity. (Since this payoff is greatly linked to the procurement cost of 3., the net cash flow, which is the sum of 3. and 4., will be less volatile than the original cash flow of 3.)

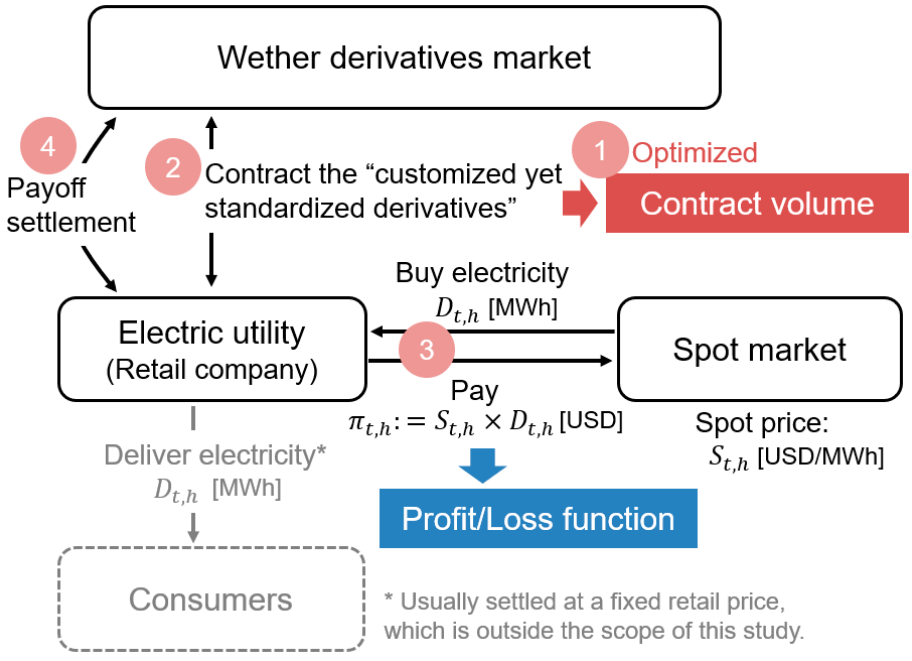

Figure A1. Transaction flowchart of "customized yet standardized derivatives".

Note that when trading "made-to-order derivatives" that cannot be fluidly traded, what the utility should optimize is the "payoff function" itself of temperature derivatives, as illustrated in "Figure 6" of [14]. In this case, the utility would contract with a specific insurer (who agreed on the optimized payoff function) rather than trading derivatives in the derivatives market. In contrast, in the case of "customized yet standardized derivatives" (and if the derivatives are traded in a fluid manner), the utility does not need to negotiate a bilateral contract with such a particular insurer; thus, the advantage of reducing trading costs may be also expected.

Appendix B. Separation of Deterministic Trends by ANOVA Decomposition

To understand ANOVA decomposition, first consider the following minimum variance hedge problem without ANOVA decomposition:

$$\underset{\Delta(\cdot)\in\mathcal{S}_{\lambda_\Delta},\ \psi(\cdot)\in\mathcal{S}_{\lambda_{\psi 1},\ \lambda_{\psi 2}}}{\text{Min}} \text{Var}[\pi_t - \Delta(t)HH_t - \psi(t,\ \varepsilon_{T,t})]. \tag{A1}$$

where $\psi(t,\ \varepsilon_{T,t})$ is the payoff function of the temperature derivative estimated as the tensor product spline function. However, since this function $\psi(t,\ \varepsilon_{T,t})$ contains a trend related to the date t, the problem that it is difficult to grasp the structure as a hedge model arises. Here, if ANOVA decomposition is applied to the bivariate tensor spline $\psi(t,\ \varepsilon_{T,t})$, the following equation can be obtained (note that in the R package "mgcv," ANOVA decomposition for the tensor product spline function can be easily calculated using the "ti" term used in the function gam() [17]):

$$\psi(t,\ \varepsilon_{T,t}) = c + \psi_t(t) + \psi_\epsilon(\varepsilon_{T,t}) + \psi_{t\epsilon}(t,\ \varepsilon_{T,t}) \tag{A2}$$

where c is a constant term, and $\psi_t(t)$, $\psi_\epsilon(\varepsilon_{T,t})$ and $\psi_{t\epsilon}(t,\ \varepsilon_{T,t})$ are obtained as zero mean functions. The univariate spline functions $\psi_t(t)$ and $\psi_\epsilon(\varepsilon_{T,t})$ are called the "main effects," which correspond to the trends that the date and temperature contribute to independently, among the original tensor product spline functions. On the other hand, the bivariate spline function $\psi_{t\epsilon}(t,\ \varepsilon_{T,t})$ is called "interaction," which corresponds to the interaction trend of the date and temperature wherein the main effect was removed from the original function. Here, the function $\widetilde{\psi}(t,\ \varepsilon_{T,t}) := \psi_\epsilon(\varepsilon_{T,t}) + \psi_{t\epsilon}(t,\ \varepsilon_{T,t})$ is the desired derivative payoff function in which just the deterministic date trend was removed from the original tensor product spline function. Additionally, by setting the contract volumes of discount bonds that depend only on t to $\widetilde{f}(t) := c + \psi_t(t)$, the minimum variance hedge problem (A1) can be modified to (3).

Appendix C. Basis Functions of the Cubic Spline and the Cyclic Cubic Spline

In this section, in order to understand the basis functions of the "cyclic cubic spline" dealt with in this study, the definition formula is briefly described along with the normal "cubic spline" (here, the contents described in [25] are summarized with supplementary explanation).

Appendix C.1. Cubic Spline Function

Cubic (regression) spline function $s_{cr}(x)$ with k knots, $x_1 \ldots x_k$ can be defined as follows by using cubic truncated power basis functions:

$$s_{cr}(x) = a_j^-(x)\beta_j + a_j^+(x)\beta_{j+1} + c_j^-(x)\delta_j + c_j^+(x)\delta_{j+1} \text{ if } x_j \le x \le x_{j+1} \tag{A3}$$

where $\beta_j = s_{cr}(x_j)$, $\delta_j = s''_{cr}(x_j)$, and the basis functions a_j^-, a_j^+, c_j^-, and c_j^+ are defined as follows:

$$\begin{aligned} a_j^-(x) &= \tfrac{x_{j+1}-x}{h_j},\ c_j^- t(x) = \tfrac{1}{6}[\tfrac{(x_{j+1}-x)^3}{h_j} - h_j(x_{j+1}-x)], \\ a_j^+(x) &= \tfrac{x-x_j}{h_j},\ c_j^+(x) = \tfrac{1}{6}[\tfrac{(x-x_j)^3}{h_j} - h_j(x-x_j)] \end{aligned} \tag{A4}$$

where $h_j = x_{j+1} - x_j$.

Here, from (A3) and (A4), this function already satisfies the condition that the value and the second derivative are equal at each knot (this fact can be proved inductively because it satisfies $s_{cr}(x_j) = \beta_j$, $s_{cr}(x_{j+1}) = \beta_{j+1}$; $s''_{cr}(x_j) = \delta_j$, $s''_{cr}(x_{j+1}) = \delta_{j+1}$). However, in order for this function to connect smoothly, the first derivative is also required to be equal at each knot. This condition can be expressed as the following matrix equation, which is derived by expanding (A3) and (A4):

$$\mathbf{B}\boldsymbol{\delta}^{-} = \mathbf{D}\boldsymbol{\beta} \tag{A5}$$

where $\boldsymbol{\delta}^{-} = (\delta_2, \ldots, \delta_{k-1})^{\top}, \delta_1 = \delta_k = 0$; **B** and **D** are defined as follows:

$$\begin{aligned} &D_{i,i} = \tfrac{1}{h_i},\ D_{i,i+1} = -\tfrac{1}{h_i} - \tfrac{1}{h_{i+1}},\ D_{i,i+2} = \tfrac{1}{h_{i+1}},\ B_{i,i} = \tfrac{h_i + h_{i+1}}{3}\ \ (i = 1 \ldots k-2); \\ &B_{i,i+1} = \tfrac{h_{i+1}}{6},\ B_{i+1,i} = \tfrac{h_{i+1}}{6} (i = 1 \ldots k-3) \end{aligned} \tag{A6}$$

Here, each element of $\boldsymbol{\delta}$ can be obtained by the matrix transformation of (A5). By substituting them into (A3) and rearranging the equation by β_i, the cubic spline function $s_{cr}(x)$ can be re-written as follows:

$$s_{cr}(x) \ = \sum_{i=1}^{k} b_i(x)\beta_i \tag{A7}$$

Appendix C.2. Cyclic Cubic Spline Function

Regarding the cyclic cubic spline function $s_{cc}(x)$, the condition that the value, first derivative and second derivative, be equal is imposed even at the start and end points of the domain (that is, knots x_1 and x_k). Even in this case, the spline can still be written in the form of (A3) and (A4), and the additional required conditions are $\beta_1 = \beta_k$, $\delta_1 = \delta_k$, and the following equations:

$$\widetilde{\mathbf{B}}\delta = \widetilde{\mathbf{D}}\beta \tag{A8}$$

where $\boldsymbol{\beta}^{\top} = (\beta_1, \ldots, \beta_{k-1})$, $\boldsymbol{\delta}^{\top} = (\delta_1, \ldots, \delta_{k-1})$; $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{D}}$ are defined as follows:

$$\begin{aligned} &\widetilde{B}_{i-1,i} = \widetilde{B}_{i,i-1} = \tfrac{h_{i-1}}{6}, \widetilde{B}_{i,i} = \tfrac{h_{i-1}+h_i}{3}, \\ &\widetilde{D}_{i-1,i} = \widetilde{D}_{i,i-1} = \tfrac{1}{h_{i-1}}, \widetilde{D}_{i,i} = -\tfrac{1}{h_{i-1}} - \tfrac{1}{h_i}\ (\ i = 2 \ldots k-1); \\ &\widetilde{B}_{1,1} = \tfrac{h_{k-1}+h_1}{3},\ \widetilde{B}_{1,k-1} = \tfrac{h_{k-1}}{6}, \widetilde{B}_{k-1,1} = \tfrac{h_{k-1}}{6}, \\ &\widetilde{D}_{1,1} = -\tfrac{1}{h_1} - \tfrac{1}{h_{k-1}}, \widetilde{D}_{1,k-1} = \tfrac{1}{h_{k-1}}, \widetilde{D}_{k-1,1} = \tfrac{1}{h_{k-1}} \end{aligned} \tag{A9}$$

Then, like the cubic spline, the cyclic cubic spline function $s_{cc}(x)$ can also be re-written as follows:

$$s_{cc}(x) \ = \sum_{i=1}^{k-1} \widetilde{b}_i(x)\beta_i \tag{A10}$$

Appendix D. Monthly RMSE and Daily Fitting Curves

Appendix D.1. Monthly RMSE

Here, the hedge errors are measured using RMSE (root mean square error). Contrary to the hedge effects, the RMSE is small for both the in-sample and out-of-sample periods in the order of demand, price, and product model, as shown in Figure A2; the annual averages of the out-of-sample period were 4.6%, 26.2%, and 33.2%, respectively. Regarding seasonality, similar shapes were confirmed for all three models; however, unlike the hedge effect, which was high in both summer and winter, the RMSE was large in winter but smaller in summer. This may be because the PJM price is prone to significant spikes in winter (while the price fluctuations in summer are relatively mild). Hence, the hedge error

is large in winter, but the correlation with temperature is strong in both summer and winter. Note that the reason the RMSE in January–February of in-sample period is significantly higher than the out-of-sample period is that in-sample data include an extreme price spike in January 2014, during which time "PJM experienced tight operational conditions and a significantly higher number of forced generator outages due to the extreme weather" [37], which is also shown in the daily price fluctuation graph in Appendix A3.

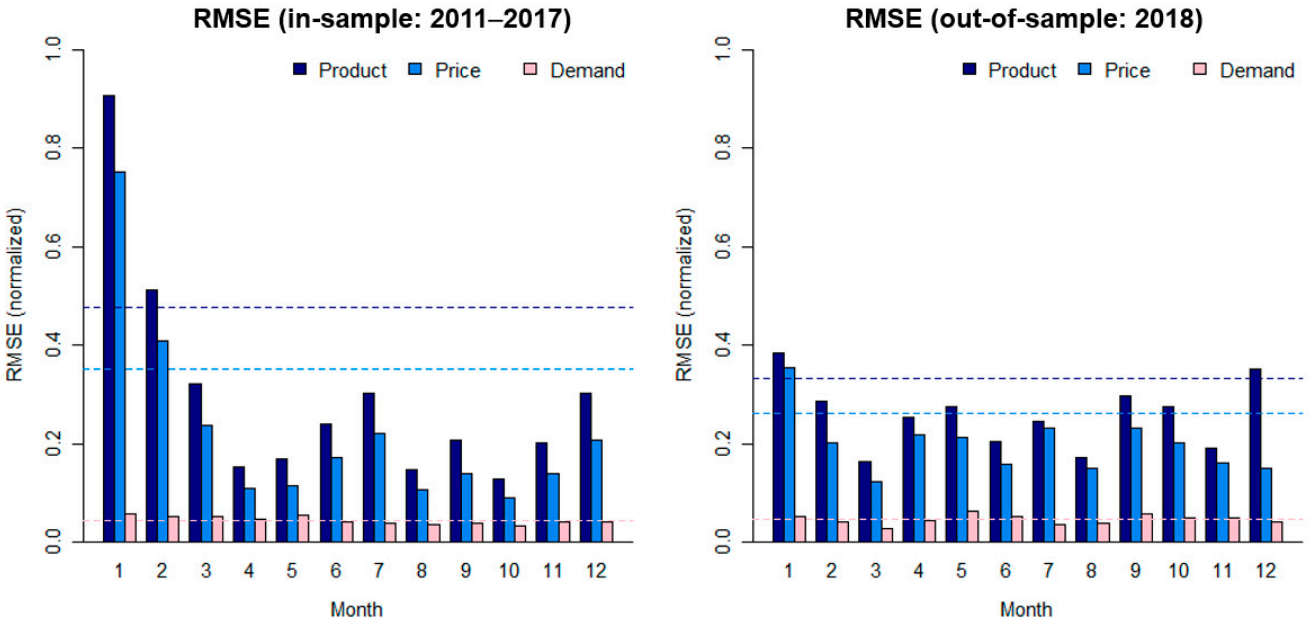

Figure A2. Monthly RMSE for each hedging target.

Appendix D.2. Daily Fitting Curves

Here, the observed value (black line) and the estimated payoff of the derivative portfolio (red line) when using all derivatives are compared for the product, price, and demand models. It can be confirmed that the derivatives' payoffs follow daily fluctuations in general, even during periods in which significant fluctuations occur, such as summer and winter.

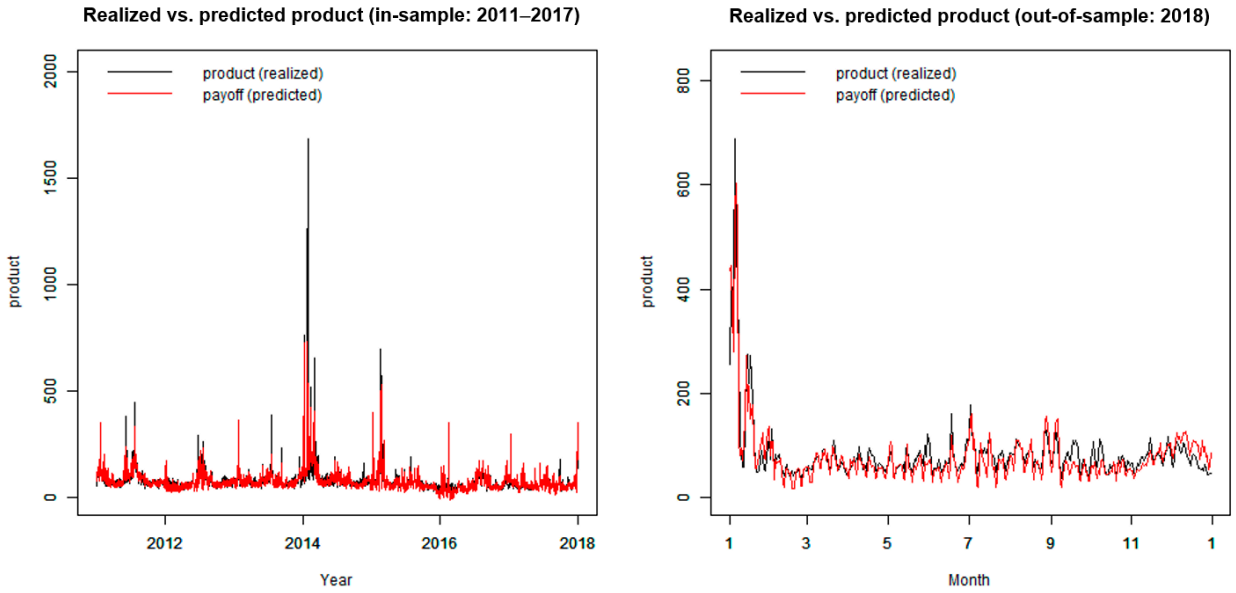

Figure A3. Comparison of realized and predicted values for the product model.

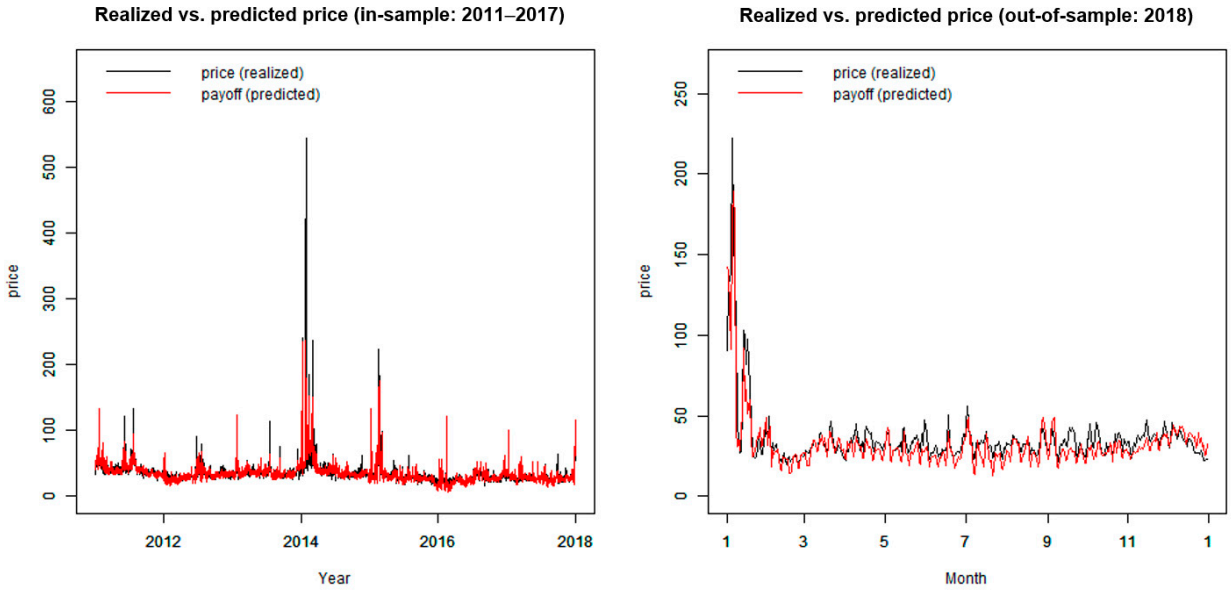

Figure A4. Comparison of realized and predicted values for the price model.

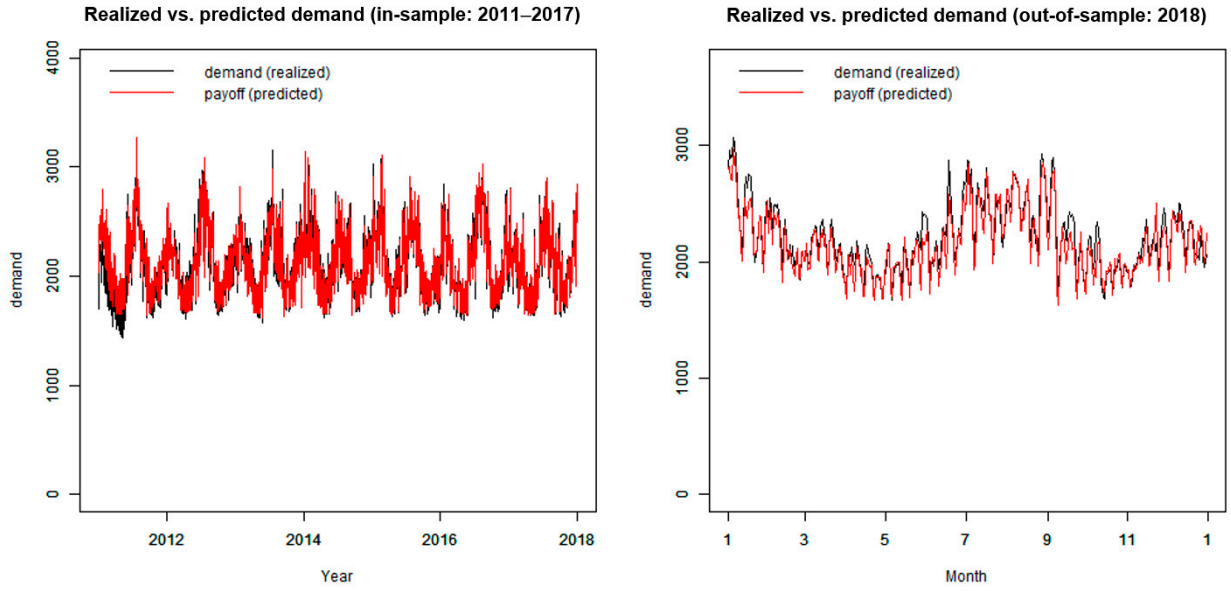

Figure A5. Comparison of realized and predicted values for the demand model.

References

1. Lee, Y.; Oren, S.S. An equilibrium pricing model for weather derivatives in a multi-commodity setting. *Energy Econ.* **2009**, *31*, 702–713. [CrossRef]
2. Lee, Y.; Oren, S.S. A multi-period equilibrium pricing model of weather derivatives. *Energy Syst.* **2010**, *1*, 3–30. [CrossRef]
3. Bhattacharya, S.; Gupta, A.; Kar, K.; Owusu, A. Risk management of renewable power producers from co-dependencies in cash flows. *Eur. J. Oper. Res.* **2020**, *283*, 1081–1093. [CrossRef]
4. Davis, M. Pricing weather derivatives by marginal value. *Quant. Financ.* **2001**, *1*, 305–308. [CrossRef]
5. Platen, E.; West, J. A Fair Pricing Approach to Weather Derivatives. *Asia Pac. Financ. Mark.* **2004**, *11*, 23–53. [CrossRef]
6. Brockett, P.L.; Wang, M.; Yang, C.; Zou, H. Portfolio Effects and Valuation of Weather Derivatives. *Financ. Rev.* **2006**, *41*, 55–76. [CrossRef]
7. Kanamura, T.; Ohashi, K. Pricing summer day options by good-deal bounds. *Energy Econ.* **2009**, *31*, 289–297. [CrossRef]
8. Yamada, Y. Valuation and hedging of weather derivatives on monthly average temperature. *J. Risk* **2007**, *10*, 101–125. [CrossRef]
9. Yamada, Y. Simultaneous optimization for wind derivatives based on prediction errors. In Proceedings of the 2008 American Control Conference, Seattle, WA, USA, 11–13 June 2008; pp. 350–355.
10. Yamada, Y. Optimal Hedging of Prediction Errors Using Prediction Errors. *Asia Pac. Financ. Mark.* **2008**, *15*, 67–95. [CrossRef]

11. Matsumoto, T.; Yamada, Y. Cross Hedging Using Prediction Error Weather Derivatives for Loss of Solar Output Prediction Errors in Electricity Market. *Asia Pac. Financ. Mark.* **2018**, *26*, 211–227. [CrossRef]
12. Yamada, Y. Simultaneous hedging of demand and price by temperature and power derivative portfolio: Consideration of periodic correlation using GAM with cross variable. In Proceedings of the 50th JAFEE Winter Conference, Tokyo, Japan, 22–23 February 2019.
13. Matsumoto, T.; Yamada, Y. Hedging strategies for solar power businesses in electricity market using weather derivatives. In Proceedings of the 2019 IEEE 2nd International Conference on Renewable Energy and Power Engineering (REPE), Toronto, ON, Canada, 2–4 November 2019; pp. 236–240.
14. Matsumoto, T.; Yamada, Y. Simultaneous hedging strategy for price and volume risks in electricity businesses using energy and weather derivatives. *Energy Econ.* **2021**, *95*, 105101. [CrossRef]
15. Hastie, T.; Tibshirani, R. *Generalized Additive Models*; Chapman & Hall: Boca Raton, FL, USA, 1990.
16. EMSC, Electricity and Gas Market Surveillance Commission. Current Status of Activation of Wholesale Power Trading, 21st Institutional Design Special Meeting Secretariat Submission Materials. 2017. Available online: https://www.emsc.meti.go.jp/activity/emsc_system/pdf/021_04_00.pdf (accessed on 11 April 2021).
17. Wood, S.N. Package mgcv v. 1.8-34. Available online: https://cran.r-project.org/web/packages/mgcv/mgcv.pdf (accessed on 11 April 2021).
18. Matsumoto, T. Forecast Based Risk Management for Electricity Trading Market. Ph.D. Thesis, University of Tsukuba, Tokyo, Japan, 2020.
19. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2009.
20. Wood, S.N. Thin plate regression splines. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2003**, *65*, 95–114. [CrossRef]
21. Eilers, P.H.C.; Marx, B.D. Flexible smoothing with B-splines and penalties. *Stat. Sci.* **1996**, *11*, 89–121. [CrossRef]
22. Eilers, P.H.; Marx, B.D. *Practical Smoothing: The Joys of P-Splines*; Cambridge University Press: Cambridge, MA, USA, 2021.
23. Wood, S.N. *Generalized Additive Models: An Introduction with R*, 2nd ed.; Chapman & Hall: Boca Raton, FL, USA, 2017.
24. Efron, B.; Stein, C. The Jackknife Estimate of Variance. *Ann. Stat.* **1981**, *9*, 586–596. [CrossRef]
25. De Boor, C. *A Practical Guide to Splines*; Springer: New York, NY, USA, 1978; Volume 27, p. 325.
26. Perperoglou, A.; Sauerbrei, W.; Abrahamowicz, M.; Schmid, M. A review of spline function procedures in R. *BMC Med. Res. Methodol.* **2019**, *19*, 1–16. [CrossRef]
27. McCulloch, J.; Ignatieva, K. Intra-day Electricity Demand and Temperature. *Energy J.* **2020**, *41*. [CrossRef]
28. Duchon, J. Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In *Constructive Theory of Functions of Several Variables*; Springer: Berlin/Heidelberg, Germany, 1977; pp. 85–100.
29. Wood, S.N.; Bravington, M.V.; Hedley, S.L. Soap film smoothing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2008**, *70*, 931–955. [CrossRef]
30. Yamada, Y.; Makimoto, N.; Takashima, R. JEPX price predictions using GAMs and estimations of volume-price functions. *JAFEE J.* **2015**, *14*, 8–39.
31. PJM Data Miner 2. Available online: http://dataminer2.pjm.com/ (accessed on 25 June 2019).
32. NOAA Climate Data Online Search. Available online: https://www.ncdc.noaa.gov/cdo-web/search (accessed on 25 June 2019).
33. EIA Henry Hub Natural Gas Spot Price. Available online: https://www.eia.gov/dnav/ng/hist/rngwhhdd.htm (accessed on 25 June 2019).
34. Thompson, T.; Webber, M.; Allen, D.T. Air quality impacts of using overnight electricity generation to charge plug-in hybrid electric vehicles for daytime use. *Environ. Res. Lett.* **2009**, *4*, 014002. [CrossRef]
35. Campbell, J.Y.; Thompson, S.B. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Rev. Financ. Stud.* **2007**, *21*, 1509–1531. [CrossRef]
36. Aguilera, A.; Aguilera-Morillo, M. Comparative study of different B-spline approaches for functional data. *Math. Comput. Model.* **2013**, *58*, 1568–1579. [CrossRef]
37. PJM. Analysis of Operational Events and Market Impacts during the January 2014 Cold Weather Events. Available online: https://www.hydro.org/wp-content/uploads/2017/08/PJM-January-2014-report.pdf (accessed on 11 April 2021).