

A Radial Basis Function Approach to Gas Storage Valuation

Alexander Boogert*, Denis Mazières†

November 13, 2011

Abstract

We introduce two new regression methods (Radial Basis Function and Tensor of Radial Basis Function) to the Least-Squares Monte Carlo (LSMC) method in the context of gas storage valuation with the spot approach. Further, we propose a new methodology which combines two existing ones in the spot approach: it combines discretisation of volume with multivariate regressions.

Our numerical examples show that these new valuation methods match and frequently manage to exceed the value of our benchmark. We also find that the valuation suffers in the case where insufficient points are placed near the volume boundaries. We expect that our new methodology in combination with the two new regression methods will prove beneficial to solve valuation problems with several volume levels in the future.

1 Introduction

Gas storages have traditionally been used to match supply and demand throughout the year. In the current environment of liberalised gas markets including third-party access to gas storage, valuation and hedging of gas storages deserves our attention. Gas storages are managed by utilities and merchants all around the world. Their motive is to either use gas storages for portfolio management purposes, pure market trading or a combination of these. In this paper we take the market-based valuation perspective, as it can create an independent benchmark. In practice three different valuation techniques are used to operate a gas storage: rolling intrinsic, spread options, and the spot approach.

The spot approach seeks to capture short-term volatility by trading (in principle) only in the spot, which is the most volatile and exhibits mean-reversion. This spot approach is the focus of the current article. The other two methods seek to capture forward volatility by trading the complete forward curve. The rolling intrinsic approach locks in the maximum value today (called the intrinsic

*EnergyQuants and Birkbeck, University of London, boogert@energyquants.com

†LCH.Clearnet and Birkbeck, University of London, denis.mazieres@lchclearnet.com

value). The holder then waits until a profitable risk-free trade can be made at a future date, which can happen when there is a switch in the forward curve. The spread options approach monetises the payoff of such switches using a spread option. For more information on the rolling intrinsic and spread options approach we refer to Boogert & De Jong [2] and the references therein. We note the latter article suggests to combine the spot approach with a static financial hedge.

At the heart of the spot approach lies an approximation of the continuation value. The continuation value is the value we attach today to having a gas storage with a specific volume tomorrow. Least-Squares Monte Carlo (LSMC) methods can be used to find such an approximation. The idea of these methods is to find a relation between state variables which are known today with the value of the contract tomorrow. First applied as a numerical technique to solve American options (see Carrière [5], Longstaff & Schwartz [13] and Tsitsiklis & Van Roy (TvR)[18]) these methods have since been applied to a variety of valuation problems. The extension to gas storage valuation can be found in Boogert & De Jong [1].

In previous literature concerning the spot approach using the LSMC method (Boogert & De Jong [1] [2], Carmona & Ludkovski [4], Denault et al. [7]) the volume aspect of a gas storage was treated in two different ways. We will introduce a new methodology which falls in-between the existing methodologies. This new methodology and its relation to the alternative methods are looked at in the next section. Part of the innovation in the recent literature is to regress on scattered data in a two-dimensional (price and volume) setting. In this article we will therefore consider how regression methods perform on scattered and grid based data.

We introduce compactly supported Radial Basis Functions (RBF) and compactly supported Tensor of Basis Functions (TBF) as an alternative regression method. RBF functions can be separated into globally supported basis functions and compactly supported basis functions. Globally supported basis functions, eg polynomials or Gaussian, are influenced by all data sites on the domain of interest; compactly supported only by data sites in the neighbourhood of their centre points. The compactly supported TBF was investigated in Hubbert & Mazières [12], and is compared in that paper against the compactly supported RBF. We will introduce RBF and TBF from a mathematical perspective in Section 3. We are interested to introduce these methods for their potential to solve high-dimensional regression problems, see Buhmann [3] and Fasshauer [8]. Our aim in this paper is to introduce these methods to gas storage valuation in a two-dimensional setting and benchmark them against a one-dimensional setting. In the current article we limit ourselves to one price and one volume dimension (and thus a two-dimensional setting). We plan to use these methods in higher dimensions in the future, eg by introducing additional volume dimensions.

To summarise, our contributions are:

- We introduce two new regression methods (RBF and TBF) which have not been used in combination with the LSMC method before in energy

finance

- We consider a new methodology to the spot approach which falls in-between the existing methodologies
- We find that these new methods provide higher, if not similar values for the gas storage when compared to our benchmark. Further, they deliver these values using fewer simulations.
- We find evidence that there needs to be sufficient points placed near the boundaries to avoid a drop in value. In our methodology we can add such points easily.

2 Relation to previous literature

In this article we consider a new methodology for the spot approach, which mixes two existing methodologies. Below we will describe these two methodologies in more detail, and highlight the differences to our methodology.

In the BdJ (Boogert & De Jong [1] [2]) methodology the volume dimension is discretised and a regression is applied to every volume level to relate prices today with the value tomorrow. Whenever a value is needed during the valuation for a volume point outside of the discrete volume levels an interpolation is used. The value is calculated for every simulation for every price-volume grid point. The alternative, to apply a single regression to both price and volume was tested by Boogert & De Jong [1], but they reported unsatisfactory convergence results for this method. For this reason we use new regression methods as discussed in Section 3.

In the CL/DSS (Carmona & Ludkovski [4], extended by Denault et al. [7]) methodology the volume dimension is not discretised and a single regression is applied to relate underlying prices today and volume tomorrow with the value tomorrow. This methodology associates one simulation to one volume level.

CL/DSS make use of two types of Least-Squares Monte Carlo (LSMC) methods. The first one, popularised by Longstaff & Schwartz [13] (or LS approach) and the other by Tsitsiklis & Van Roy [18] (or TvR approach). The first one determines the value from subsequent cash flows, while the second one determines the value by reading it from the regression surface. For more detail we refer to the discussion in Glasserman [9]. As noted by Carmona & Ludkovski the TvR approach has less variance but more bias compared to the LS approach when applied to price American options.

The CL/DSS methodology (called BLSM by Carmona & Ludkovski) aims to follow as much as possible the LS approach, and only use occasionally the TvR in specific circumstances. This comes at a computational cost as the LS approach is slower than the TvR approach. In order to follow the LS approach CL/DSS employ a special procedure to determine how to scatter the price-volume data sites. The procedure relies on a mechanism to forecast relevant volume levels in a backward fashion, which aims to maintain ‘forward optimal paths’. When

CL/DSS cannot maintain forward optimal paths, they randomise the volume level and apply the TvR approach instead.

In our method we take an intermediate approach to the distribution of price-volume data sites. We will develop on this later. Further, we refrain from using the procedure described above. We simply use a standard discretisation or in some instances randomisation of the volume level. One of the reasons we do not use the mechanism to forecast relevant volume levels is that we expect our regression to experience boundary effects, which is a concentration of the error on the boundaries of the domain. With such a mechanism automatically defining data sites, we would lose the control to allocate sufficient data sites on the boundaries. This would obviously deteriorate the accuracy in case a regression is carried out with too few data sites in the boundary regions. The second reason for not using this mechanism is due to both operational restrictions and the shape of the forward curve: we can be wrong in our forecast and the approach will place the volume points incorrectly. As a consequence of not using this mechanism we are now free to set a grid according to other criteria. For example, if we were to use an adaptive grid we would gain on computational speed (this is especially true in TBF setting, see the explanation on TBF below). The price we pay is that we always use the TvR approach, which is not as accurate as the LS method. We expect our improved regression methods to balance against this difference.

RBF are applied in high dimensional interpolation and regression, eg in image processing. In the past decade RBF has been a very active domain of research, important references are Fasshauer [8], Schaback [15] and Wendland [19]. In finance, the main application has been to solve PDE. An application of globally supported RBF in LSMC has been shown by Grau [10]. He uses sparse polynomial basis functions to value moving window Asian options, callable convertible bonds, and thin plate basis functions to value a barrier option. As far as we know compactly supported RBF have not been applied in energy finance before. When different basis functions are compared, in the context of LSMC, it is often found that power functions work well. This was concluded by Longstaff & Schwartz [13], Moreno & Navas [14] and Stentoft [17] for a variety of financial options and Boogert & De Jong [1] and [2] for gas storage.

Carmona & Ludkovski [4] and Denault et al. [7] include volume into the regression for gas storage. Further, Carmona & Ludkovski [4, p. 367] indicate they experimented with exponential and polynomials as basis functions, but do not provide details as to which exact bases are applied and no comparison is provided. Denault et al. [7, p.8] indicate they experimented with different bases (stepwise, piecewise linear and polynomials) and concluded simple powers work well, but no comparison is provided either. In this article we present a comparison between compactly supported RBF and simple powers.

3 Two approaches for multivariate regressions

We are now going to present the tools, ie kernels and regressions, that we will apply later in the gas storage valuation. Two types of kernel will be introduced. First, the Radial Basis Function and then the Tensor of Basis Functions.

3.1 Radial Basis Function

Numerical methods based on Radial Basis Functions (RBF) are well known for their dimension “blindness”. For this reason they are in theory ideal candidates for solving high dimensional problems. This dimension blindness (and the origin of the name radial) comes from the fact that their support is defined by the Euclidean norm $\|\cdot\|$, or put simply by the distance between two points defining the support of the basis. The support is radial around a so called centre point.

Another fundamental feature of RBF is their strictly positive definiteness. So that when used in the context of interpolation or regression, they lead to positive definite matrices which are then invertible. This is a key property which we will implicitly make use of throughout this work.

3.1.1 Choosing a compactly supported basis function

RBF comes in different flavours. Some are globally supported like the Gaussian, the thin plate spline and the generalised multiquadric. Others are compactly supported, such as the family of Wendland and Wu functions, see [8] and [19]. All work in a similar manner as basis functions for interpolation and regression.

Among the compactly supported, most have an additional feature which scales their support. This is done through the adjustment of the shape factor ϵ . An increase in ϵ shrinks the support of the compactly supported basis functions, while a decrease in ϵ broadens the support.

Throughout this paper we will use a specific type of compactly supported basis function, from the family of Wendland functions. The one we have chosen is the Wendland C^2 defined in \mathbf{R}^3 . This means that this function has a smoothness of order 2 and can be used in problems up to and including three dimensions. Similar functions are available for higher or lower orders of smoothness and are available for higher dimensional problems. The equation for this function is:

$$\phi(r) = (1 - \epsilon r)_+^4 (4\epsilon r + 1) \tag{1}$$

We provide an illustration of our basis function in Figure 1. In the top diagram we show the impact of the shape factor on the one-dimensional Wendland function. In this case the centre point is set at 0. On first sight these functions look like Gaussians, but they contain a definite cutoff point after which the value is zero, hence their compact support nature.

We selected the Wendland C^2 , see Equation (1), as it is sufficiently smooth for our problem but not too smooth. With a higher order of smoothness for the

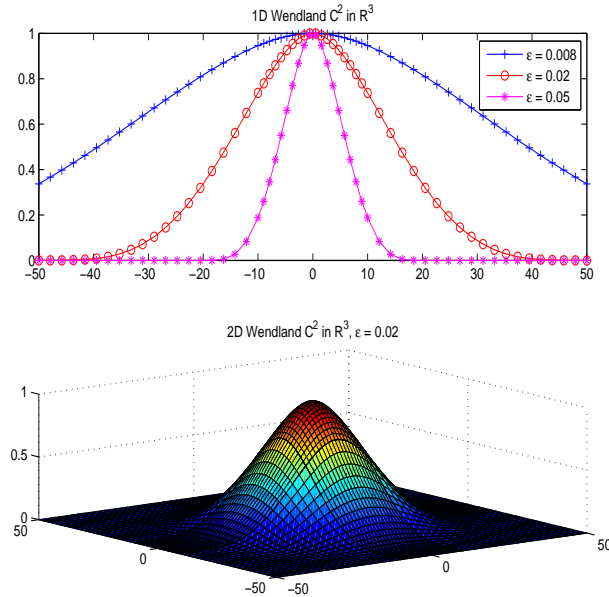


Figure 1: Impact of shape factor ϵ on support for Wendland C^2 defined in \mathbf{R}^3 in 1 and 2 dimensions

basis function we would be at risk to over-fit the problem. The family of Wendland functions offers a polynomial order of smoothness. Globally supported basis functions, like the Gaussian, are infinitely smooth. Although this sounds like an attractive feature, it leads to interpolation and regressor matrices with large condition numbers, which are harder to invert. This backs up our choice for the compactly supported basis function.

3.1.2 RBF based interpolation

We have chosen to first introduce RBF based interpolation and then move to the RBF based regression, which we will use further in this article. There are two reasons for this. First, in the RBF literature we observe it is mostly written in the context of interpolation, see Hubbert & Mazières [12] and Wendland [19]. Second, the exposition is easier starting with the interpolation approach rather than straight with the regression one.

The d -dimensional scattered data interpolation problem is posed as follows: let $\Omega \subset \mathbf{R}^d$ denote the domain of a d -dimensional function Y . Suppose that the only values we have for Y are those at a set of N distinct locations $\{\mathbf{x}_i\}_{i=1}^N \subset \Omega$, which we call data sites. The problem is to find an interpolant $s_f : \Omega \rightarrow \mathbf{R}$ such that the value of the function s_f matches the value Y in all data sites \mathbf{x}_i :

$$s_f(\mathbf{x}_i) = Y(\mathbf{x}_i) \quad i = 1, \dots, N. \quad (2)$$

Let's consider s_f as an expansion of RBF such as,

$$s_f(\mathbf{x}_i) = \sum_{j=1}^M \alpha_j \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) \quad (3)$$

where α are the interpolation coefficients and ϕ are the RBF applied to the centre points \mathbf{x}_j , $j = 1 \dots M$ and associated with M basis functions. Note that these centre points usually do not coincide with the data sites. Equation (2) can be rewritten with Equation (3),

$$Y(\mathbf{x}_i) = \sum_{j=1}^M \alpha_j \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) \quad i = 1, \dots, N \quad (4)$$

In matrix form this is,

$$\mathbf{Y} = \Phi \alpha \quad (5)$$

where Y is the vector of N values and α is the vector of M interpolation coefficients, while Φ is the N by M square interpolation matrix (for interpolation, $M = N$). This interpolation matrix results from applying the RBF to the distance matrix, where this distance matrix is the Euclidean norm of the data sites taken against the centre points,

$$\text{distance matrix} = \|\mathbf{x}_i - \mathbf{x}_j\| \quad i = 1, \dots, N, \quad j = 1, \dots, M$$

3.1.3 RBF based regression

The regression within a RBF setting is a mere extension of the interpolation presented above and is effectively an over-determined case where $N > M$. Now, let's call β the regression coefficients, previously known as α in interpolation, and write Equation (4) and (5) in the regression setting.

$$Y(\mathbf{x}_i) = \sum_{j=1}^M \beta_j \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) \quad i = 1, \dots, N \quad (6)$$

In matrix form this is,

$$\mathbf{Y} = \Phi \beta \quad (7)$$

where Y is the vector of N values and β is the vector of M regression coefficients, while Φ is the N by M regressor matrix but is not square any more as now $N > M$. Further, if the regressor matrix Φ is generated from a set of distinct centre points, it is full rank M . Hence $\Phi' \Phi$ is non-singular and invertible. This derives from the strictly positive definiteness property of the RBF, which we defined earlier in section 3.1.

Since Φ is full column rank, the solution of equation (7) can be obtained as

$$\beta = (\Phi' \Phi)^{-1} \Phi' Y \quad (8)$$

As it was said before, the matrix $\Phi'\Phi$ is in theory strictly positive definite, hence invertible. However in practice this matrix happens to be, very often, poorly conditioned and close to singularity. This comes from the fact that a fixed support is used for the basis function (RBF), which we use to construct the regressor matrix Φ . The construction of the regressor cross-product matrix $\Phi'\Phi$ worsens the condition number found for the regressor matrix Φ . In addition, the regressor matrix can end up with a high condition number, when datasites almost overlap. It is possible to compute β using Equation (8), but it has to be in combination with the computation of optimal shape factors (which scale the support of the basis) at every time step. We will discuss this in more detail in Section 5.3. We experimented with this, but found that this is very time consuming. A computationally cheaper alternative is to directly solve the linear system of Equation (7) with a linear system solver or to use the Moore-Penrose inverse on the regressor matrix Φ . We experimented with both and found that they give very similar, if not identical, results.

3.2 Tensor of Basis Function

3.2.1 Concept

An alternative to the RBF is to consider a tensor of RBF. The idea behind the Tensor of Basis Functions (TBF) using univariate Wendland functions was introduced in Hubbert & Mazières [12]. Despite the TBF only using univariate basis functions, it can be used for multivariate interpolations or regressions, ie for any number of dimensions. This idea stems from the fact that the product of strictly positive definite basis functions is strictly positive definite. This is effectively the Schur product Theorem in disguise, see Cheney & Light [6] or Horn [11] for a formal definition. Further, despite obvious structural differences the TBF manages to keep the main properties of the RBF, which are the dimension blindness and the strictly positive definitiveness mentioned earlier.

3.2.2 TBF based interpolation

For the same reasons we introduced the RBF in the interpolation setting first. Let's start to introduce the TBF for the interpolation problem too. This will make the transition to the TBF based regression a lot easier. This is as most of the theory around tensor of basis functions focuses on interpolation, see Hubbert & Mazières [12] and Cheney & Light [6]. We keep here the same interpolation problem presented in the previous Section in Equation (2). Here the value $Y(\mathbf{x}_i)$ living on a d -dimensional domain will be interpolated with an expansion of TBF, namely by a product of univariate basis functions. This expansion is:

$$s_f(\mathbf{x}_i) = \sum_{j=1}^M \alpha_j \prod_{l=1}^d \phi^1(\|x_i^l - x_j^l\|)$$

$$= \sum_{j=1}^M \alpha_j \phi^1(\|x_i^1 - x_j^1\|) \phi^2(\|x_i^2 - x_j^2\|) \dots \phi^1(\|x_i^1 - x_j^1\|) \dots \phi^d(\|x_i^d - x_j^d\|)$$

All the x^l are the l -coordinates of the data sites $\mathbf{x}_i = [x_i^1, x_i^2, \dots, x_i^1, \dots, x_i^d]$, $i = 1, \dots, N$. α_j are the regression coefficients. $\phi^1(\cdot), \phi^2(\cdot), \dots, \phi^d(\cdot)$ are the TBF applied to the data sites \mathbf{x}_i and associated with M centre points \mathbf{x}_j .

It is useful to see the TBF defined in d -dimensional space. However, since in the next section we will limit ourselves to two dimensions, we will revert to a more friendly notation denoting by p dimension 1 and by v dimension 2.

$$s_f(\mathbf{x}_i) = \sum_{j=1}^M \alpha_j \phi_p(\|p_i - p_j\|) \phi_v(\|v_i - v_j\|) \quad i = 1, \dots, N \quad (9)$$

where α_j are the regression coefficients and $\phi_p(\cdot)\phi_v(\cdot)$ are the TBF applied to the data sites $\mathbf{x}_i = [p_i, v_i]$, $i = 1, \dots, N$ and associated with M centre points.

On first sight when we compare the TBF to the RBF we note that only the construction of the basis is different, but this alone has profound consequences on the performance and stability of the method. With TBF, as opposed to with RBF, each basis works independently in specifically designated dimensions. We will see later that we can use this design feature to our advantage, when we will consider compactly supported basis functions equipped with shape factors.

Now let's rewrite the interpolation problem set in Equation (2) with Equation (9),

$$Y(\mathbf{x}_i) = \sum_{j=1}^M \alpha_j \phi_p(\|p_i - p_j\|) \phi_v(\|v_i - v_j\|) \quad i = 1, \dots, N$$

In matrix form this is,

$$\mathbf{Y} = \Phi_p \otimes \Phi_v \alpha$$

where Y remains the vector of N values and α is again the vector of M interpolation coefficients, while $\Phi_p \otimes \Phi_v$ is the N by M square interpolation matrix (for $N = M$). The operator \otimes is the tensor product as well known as the Kronecker product, see Cheney & Light [6] for more details.

3.2.3 TBF based regression

Like for the RBF one can re-interpret the interpolation with TBF, seen above, into a regression setting. This is written,

$$Y(\mathbf{x}_i) = \sum_{j=1}^M \beta_j \phi_p(\|p_i - p_j\|) \phi_v(\|v_i - v_j\|) \quad i = 1, \dots, N \quad (10)$$

where $N > M$

In matrix form this is,

$$\mathbf{Y} = \Phi_p \otimes \Phi_v \beta$$

Then we can compute β such as,

$$\beta = \left[(\Phi_p \otimes \Phi_v)' (\Phi_p \otimes \Phi_v) \right]^{-1} (\Phi_p \otimes \Phi_v)' Y \quad (11)$$

We follow here the method introduced by Hubbert & Mazières [12] to inverse the TBF. This is done by using properties of tensor algebra which allow rewriting Equation (11) into

$$\beta = \left(\Phi_p' \Phi_p \right)^{-1} \Phi_p' \otimes \left(\Phi_v' \Phi_v \right)^{-1} \Phi_v' Y \quad (12)$$

The computational speed-up is achieved by applying the Moore-Penrose algorithm separately to the two matrices $\Phi_p' \Phi_p$ and $\Phi_v' \Phi_v$ which need to be inverted, rather than to the larger matrix $(\Phi_p \otimes \Phi_v)' (\Phi_p \otimes \Phi_v)$.

4 Gas storage valuation

In this Section we provide a mathematical description of our new methodology of the spot approach for gas storage valuation. We refer to [1] and [4] for an extended description of the gas storage problem. We will illustrate the univariate and new multivariate regressions using RBF and TBF in Section 4.1 and 4.2 below.

4.1 Spot approach using univariate regression

We denote by $v(t)$ the gas volume in storage at the start of day t , and by Δv the action during day t . We denote by continuation value $C(t, S(t), v(t+1))$ the value we attach today to having a gas storage with volume $v(t+1)$ tomorrow given a spot price today of $S(t)$. We get to this volume after taking an allowed action Δv starting at volume $v(t)$ today. If we denote the set of allowed actions by $\mathcal{D}(t, v(t))$ and the immediate payoff of an action by $h(S(t), \Delta v)$, then our pricing problem can be defined as finding for every combination $(S(t), v(t))$ the optimal action as defined by

$$\arg \max_{\Delta v \in \mathcal{D}(t, v(t))} \{h(S(t), \Delta v) + C(t, S(t), v(t+1))\} \quad (13)$$

We will discretise the volume into n volume levels. For our benchmark we will regress the state variable spot price on future value Y at volume level $v(t+1; n)$. If we have m simulations and n volume levels, this means we regress m future values at volume level $v(t+1; n)$ on m spot prices and we do this for every volume level. Thus in total we perform n regressions per time step. Our benchmark Cubic1D regression method is the following regression

$$Y = \beta_0 + \beta_1 S(t) + \beta_2 S^2(t) + \beta_3 S^3(t) \quad (14)$$

Using the regression coefficients, we can find estimated future values for each future volume level. If an action Δv takes us in-between two future volume levels we use linear interpolation between the two estimated future values.

In the first part of our numerical example in Section 5 we will work with an equidistant volume grid. In the second part of our numerical example we will show how a non-equidistant volume grid can be used. In the first part of our numerical example we will provide a comparison between univariate and multivariate regression methods. On the univariate side, we will consider the cubic mentioned above in Equation (14) and the Wendland C^2 RBF described earlier in Section 3.1, but this time as a univariate.

On the multivariate side, regression methods have been presented in the previous Section. Next we will illustrate how to use them for gas storage valuation.

4.2 Spot approach using multivariate regression

We re-introduce here the multivariate regression methods introduced in Section 3, but now in a gas storage setting. This means we will provide a specific gas storage interpretation to the previous general description. We will keep the same notation as in Section 4.1. In a gas storage setting, the RBF regression Equation (6) becomes

$$Y(\mathbf{x}_i) = \sum_{j=1}^M \beta_j \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) \quad i = 1, \dots, N \quad (15)$$

where Y is the future gas storage contract value and every data site $\mathbf{x}_i = [p_i \ v_i]$ contains the coordinates of a *price-volume* data site chosen on the price-volume surface prevailing at every time step. The distribution of the price-volume data sites can be either a well ordered grid or scattered data.

In gas storage setting the TBF regression Equation (10) becomes

$$Y(\mathbf{x}_i) = \sum_{j=1}^M \beta_j \phi_p(\|p_i - p_j\|) \phi_v(\|v_i - v_j\|) \quad i = 1, \dots, N \quad (16)$$

where p_i and v_i are the projections of \mathbf{x}_i data sites on respectively the price and volume dimensions. Further, β_j are the regression coefficients and $\phi_p(\cdot)$ and $\phi_v(\cdot)$ are the bases of the TBF applied to the data sites $\mathbf{x}_j = [p_j \ v_j]$, $j = 1, \dots, M$ and associated with M centre points. The basis $\phi_p(\cdot)$ works only in the price dimension and $\phi_v(\cdot)$ works only in the volume dimension. Because the TBF works on the projections of the data sites rather than on the data sites themselves, the method benefits greatly from using data on a grid.

5 Numerical example

5.1 Example setting

In our numerical example we study three different storages with characteristics given in Table 1. In addition to these characteristics we set the minimum, start and end volume equal to zero. We assume there are no costs involved other than those related to buying and selling gas and set the interest rate to zero. The trading period is one year. The cases represent an oil reservoir (case 1), a fast (case 2) and finally a very fast (case 3) salt cavern. Case 3 represents a case where the maximum volume is not an exact multiple of the injection and withdrawal rate. This allows us to study the effect of interpolation between volume levels.

	case 1	case 2	case 3
Max. injection rate	1	2	2.7
Max. withdrawal rate	1	5	5.8
Max. volume	100	100	50

Table 1: Characteristics of the three different storage cases in our numerical example

In this article we take as a price process the mean-reverting one-factor Schwartz [16] model, also referred to as a discrete-time Ornstein-Uhlenbeck process. In log-terms it is given by

$$d \ln P^{ST}(t) = \kappa \left[\mu(t) - \ln P^{ST}(t) - \frac{(\sigma^{ST})^2}{2\kappa} \right] dt + \sigma^{ST} dW^{ST}(t) \quad (17)$$

where $P^{ST}(t)$ is the spot price and the mean level $\mu(t)$ is a deterministically time varying function. The daily mean-reversion rate κ and volatility σ^{ST} are assumed to be constant. In our experiments we use $\sigma^{ST} = 150\%$ and $\kappa = 12\%$ (annualised). We set $\mu(t)$ equal to the forward curve as shown in Figure 2. Whenever we present a gas storage valuation, this is based on six different seeds.

5.2 Comparison 1D with 2D regression methods

In our numerical example we compare the valuation coming from four different regression methods. We use 500 simulations and 100 volume levels, and compare two 1-dimensional regression methods (Cubic1D, RBF1D) with two 2-dimensional regression methods (RBF2D, TBF2D). The Cubic1D is our benchmark method. It uses the regression in Equation (14) in the price dimension and interpolates the volume dimension. In RBF1D we replace the cubic regression by a one-dimensional RBF regression on prices. The 2-dimensional regression methods interpolate over price and volume as in Equation (15) and

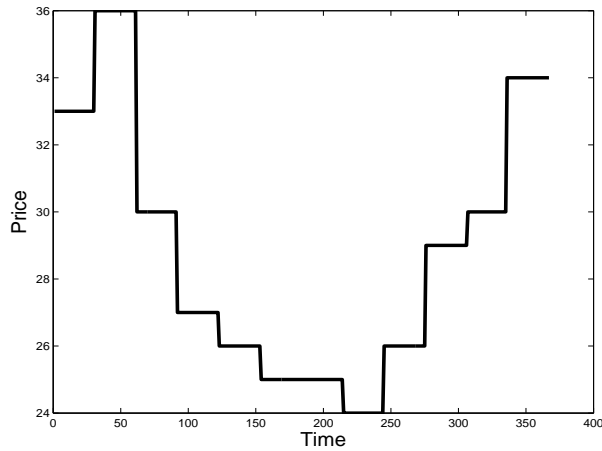


Figure 2: Initial forward curve

		Cubic1D	RBF1D	RBF2D	TBF2D
case 1	mean	1072.8	1073.4	1074.7	1073.5
	stdev	0.4	0.2	0.2	0.2
case 2	mean	1378.9	1383.1	1378.3	1379.3
	stdev	2.4	1.4	1.2	1.5
case 3	mean	1466.1	1471.4	1467.6	1471.4
	stdev	3	1.6	1.7	1.8

Table 2: Valuation of the three storage cases using the four different regression methods, 500 simulations

(16). We provide the valuation results for the four regression methods and the three storage cases in Table 2.

We conclude that all methods provide a similar value: all methods deliver a value within 1% of each other. This is in line with the results provided by Carmona & Ludkovski [4] and Denault et al. [7], who present results within 2% of their benchmark value (coming from an alternative method). It is however important to note that the new methods we proposed (RBF1D, RBF2D and TBF2D) consistently outperform our benchmark Cubic1D.

In Table 2 we also see that Cubic1D provides less stable valuations for the same number of simulations; it has a higher standard deviation of the mean results than the alternative methods. An alternative way to compare the stability of a simulation algorithm is to compare in-sample and out-of-sample valuations, see e.g. Boogert & De Jong [1]. An in-sample valuation is the value resulting from implementing the learned decision rules on the same set of simulations. This value is known to have an upper bias in the original TvR approach due to Jensen's inequality, see Glasserman [9]. An out-of-sample valuation is the value resulting from implementing the learned decision rules on a new set of simula-

tions, and has a lower bias. We present the results in Table 3. We confirm that the in-sample valuation has an upper bias: the out-of-sample valuation adjusts the value downwards. In Table 2, and the remainder of the article, we report out-of-sample values.¹ We also observe that the 1D regression methods have a higher difference between in-sample and out-of-sample than the 2D regression methods, especially Cubic1D. This is a good point for the 2D methods.

		Cubic1D	RBF1D	RBF2D	TBF2D
case 1	in sample	1073.7	1073.7	1075.0	1073.7
	out of sample	1072.8	1073.4	1074.7	1073.5
	difference	-0.9	-0.3	-0.4	-0.2
case 2	in sample	1385.3	1385.9	1379.5	1380.4
	out of sample	1378.9	1383.1	1378.3	1379.3
	difference	-6.3	-2.8	-1.2	-1.1
case 3	in sample	1473.8	1475	1469.1	1472.6
	out of sample	1466.1	1471.4	1467.6	1471.4
	difference	-7.7	-3.6	-1.5	-1.2

Table 3: Comparison in and out-of-sample valuation of the three storage cases, 500 simulations

Next, we consider the impact of the number of simulations on the valuation. We ran the valuations of the three storage cases using the four regression methods for different numbers of simulations (50, 100, 150, 250 and 500). The results are shown in Figure 3. In Table 2 we showed how, for 500 simulations, all valuations came out in a tight range. In Figure 3 we note that TBF2D and RBF1D reach already their maximum value for 100 simulations, whereas RBF2D settles at 250 simulations and Cubic1D increases until 500 simulations. This is in line with the results in Table 2: Cubic1D has a higher standard deviation of the mean results. This is as well true for a lower number of simulations.

From a performance perspective we have found TBF2D to be significantly faster than RBF2D by several folds. This is thanks to the inversion method used for the TBF2D method, see Equation (12). This method was initially introduced by Hubbert & Mazières [12], and we see here a similar computational speed improvement as they did. Their results indicate the TBF is 30 (2 dimensions), 80 (3 dimensions) and 680 (4 dimensions) times faster than the RBF method. This hints that TBF should be the method to favour should we consider to carry out regressions in higher dimensions.

Therefore we conclude that the 2D regression methods are more stable than their 1D counterparts. In particular TBF2D appears promising given its good performance for a low number of simulations. This could be due to the fact that there are two free shape factors in TBF2D. We will discuss the impact of the shape factors on the valuation below, where we will also introduce the way these are determined.

¹We assume that the MiTvR valuation provided by Carmona & Ludkovski [4] are in-sample values: they are higher than the alternative.

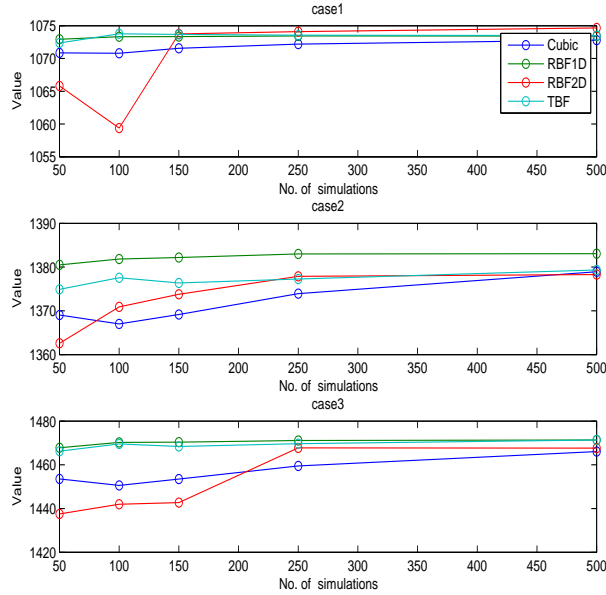


Figure 3: Impact of the number of simulations on the valuation using the four different regression methods for the three storage cases: case1, 2 and 3 respectively from top to bottom

5.3 Impact of shape factor

5.3.1 Fixed shape factor for RBF

One of the main features of compactly supported RBF and TBF is the shape factor ϵ . We introduced the notion of shape factor in Section 3.1. From Section 4.2 we know that RBF2D has a single shape factor ϵ_{RBF} , whereas TBF2D has two shape factors ϵ_p and ϵ_v . In Figure 1 we showed the impact of changing ϵ on the support of the Wendland function, which we use as basis function. Its formula was given in Equation (1). In general, a large ϵ contracts the support, while a small ϵ broadens the support of the compactly supported basis function.

In the gas storage problem we note that the support changes over time. This is especially relevant for the volume dimension: the possible volume range increases from a single point (at the start of the contract) to the maximum allowed interval (in the middle of the contract) and back to a single point again at the end of the contract. This makes it difficult for fixed shape parameters to perform well throughout the whole trading period.

Within RBF2D a single shape factor has to cater for the changing support in both the volume and price dimension. For this reason it seems hard to adjust ϵ_{RBF} over time. Fortunately, RBF2D based valuations are very indulgent concerning the choice of the shape factor, as seen in Figure 4. We provide in the Figure 4 the value resulting from different choices of ϵ_{RBF} for the three storage

cases. It is obvious from these graphs that the RBF2D valuation reaches its maximum value for a broad range of $\epsilon \approx [10^{-5}, 10^{-4}]$ and that this range can be used for different storage cases. Hence, we are pretty safe picking one ϵ_{RBF} for all the valuations we require.

If we apply a similar trial and error process to determine a correct value for the ϵ_p and ϵ_v , we find that the TBF2D method is more sensitive than the RBF2D to a small change in either shape factor. As a consequence the TBF2D based valuation method has more difficulties to converge to the correct contract value using fixed shape factors. We therefore seek an adaptive method to set the two shape factors over the trading period. This is discussed next.

5.3.2 Adaptive shape factor for TBF

As far as we know there is no exact formula available to determine the optimal ϵ_p and ϵ_v . In order to approximate these shape factors we will first need to introduce the concept of projection of fill distance on the different dimensions of the problem, see Hubbert & Mazières [12] for details. In the context of gas storage the fill distance is a measure of the data distribution in each dimension of the problem, that is price and volume. The fill distances, h_p and h_v for respectively the price and volume dimension are defined as,

$$h_p = \sup_{p \in \Omega_p} \min_{p_j \in \chi_p} \|p - p_j\| \quad (18)$$

$$h_v = \sup_{v \in \Omega_v} \min_{v_j \in \chi_v} \|v - v_j\| \quad (19)$$

The fill distance h_p (respectively h_v) indicates how well the data in the set χ_p (respectively χ_v) fills out the domain Ω_p (respectively Ω_v). A geometric interpretation of the fill distance is given by the radius of the largest possible empty ball that can be placed among the data locations inside the domain. Finding the optimal shape factor is an active domain of research in RBF, but to our knowledge no exact formula is available for it. Nevertheless, we know that there is an inverse relationship between the fill distance and the optimal shape factors in the RBF setting, see Fasshauer [8] and Schaback [15]. In the literature this is known as stationary setting or adaptive method. Here we apply this method to the TBF setting in order to approximate ϵ_p and ϵ_v . We define the shape factors as,

$$\epsilon_p = \frac{c_p}{h_p} \quad (20)$$

$$\epsilon_v = \frac{c_v}{h_v} \quad (21)$$

where c_p and c_v are constants, also known as fixed base scale factors.

When we experimented with the adaptive formulas in Equation (20) and (21) for a variety of cases, we found a pair of fixed based scale factors (c_p, c_v) that

work well for all of them. We investigate this further using Figure 5 where we show the impact of changing the fixed based scale factors for the three different storage cases. We see in Figure 5 that the c_v has very little impact on the value, except when c_v gets bigger than 0.01, where the contract value becomes erratic. On the other hand, c_p has a significant impact on the value. Furthermore, it seems that in most cases the same c_p provides the highest value for the contract despite working with different types of storages.

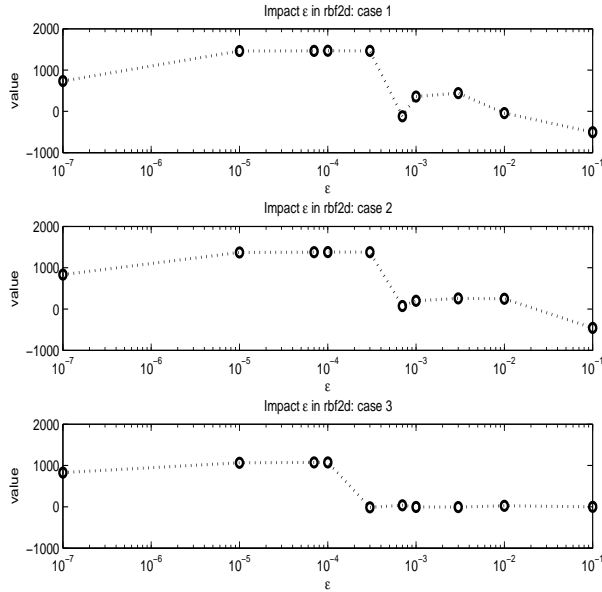


Figure 4: Impact of epsilon on the value using RBF2D regression

5.4 Removing data sites

We are aware that high dimensional regressions are computationally expensive. Hence to make our methods future proof we investigate the idea of removing data sites from the full-grid we have been using so far in our regression. Ultimately, this is to reduce the size of the regressor matrices and speed up the computation. The additional benefit of reducing the size of these matrices is that it tends to lower their condition numbers, which in turn helps solving the linear systems of regression equations. This removal process brings us from a full-grid data set to a scattered data set. In the context of gas storage this takes us from a regression on a full-grid to a regression on scattered data sites, as in Carmona & Ludkovki [4] and Denault et al. [7].

In our experiment we start from the full-grid approach and take out half of the data sites. We compare different ways to take them out. First, we take out data sites completely at random. Next, data sites are removed randomly while

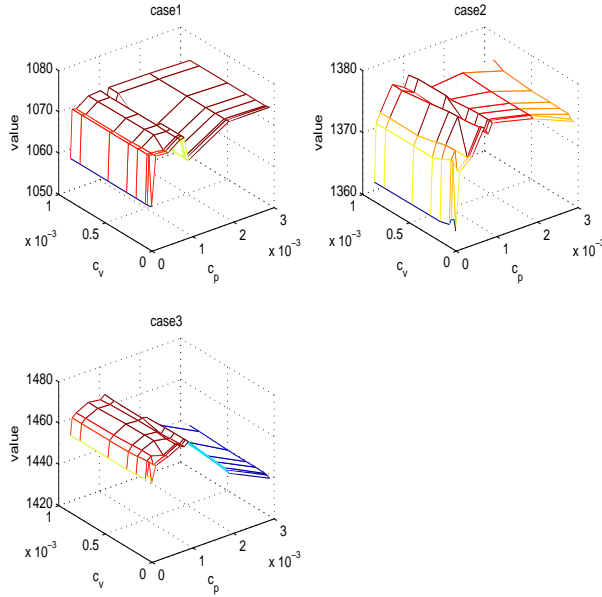


Figure 5: Impact of fixed base scale factors on the value using TBF2D regression

p layers of data sites are protected on the volume level sides; these data sites are never taken out. This approach is illustrated in Figure 6 where we protected 5 layers of data sites (or $p = 5$). In this experiment we remove 50% of the data sites, and consider different levels of protection $p = 0, \dots, 5$ for the volume level boundaries.

Results are provided in Figures 7 for RBF2D and Figure 8 for TBF2D. If we compare the valuation for 150 and 75 simulations in Figure 3, we would expect to find a similar difference in value in this current experiment. This is because taking out 50% of the data sites compares to reducing the number of simulations from 150 to 75 whilst keeping the number of volume levels constant at 100. Effectively we start with a full grid with 15,000 data sites and drop to 7,500 scattered data sites.

However, a clear drop in value is observed when no protection ($p = 0$) is applied against the removal of data sites from the volume level boundaries. But as soon as a few layers are protected, the gas storage value regains its correct value. On the other hand if too many data sites are protected on the sides, fewer data sites remain in the middle and the quality of the regression suffers. We can observe this phenomena in both the RBF2D and TBF2D approach, see Figure 7 and 8. This confirms that both these methods are very sensitive to the presence of sufficient support on the boundaries of the volume domain of the regression. We have not observed the same phenomenon in the price dimension.

A potential explanation for the success of the protection method of the boundary layer can be found in the context of interpolation and regression: the

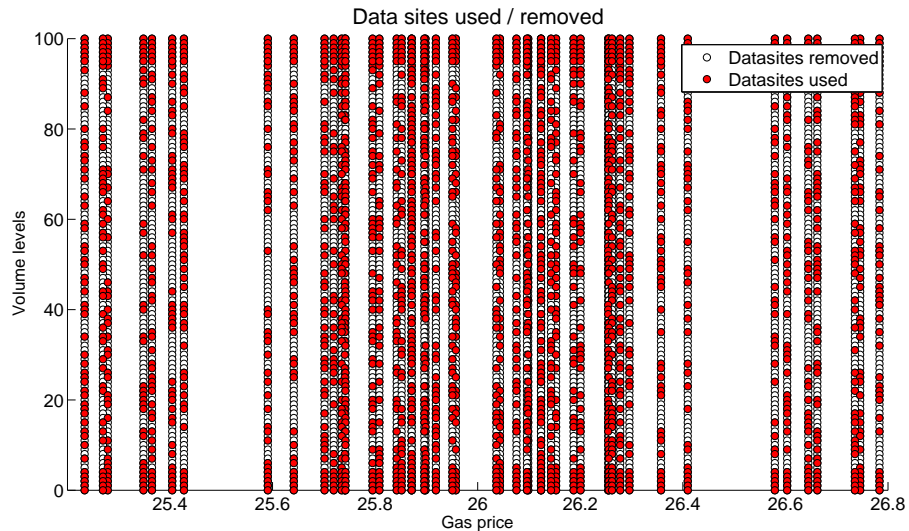


Figure 6: Illustration of volume layer protection in a grid. In this example 5 volume layers of data sites are always preserved.

highest error values tend to cluster on the boundaries of the domain. One way to cure this problem is to concentrate sufficient data sites on the boundaries of the domain of interest. This is a fairly crude approach, but it has been reported to work well in Fasshauer [8] in the context of solving a PDE with Kansa's method. Despite this observation, we are surprised that this effect is so important in our experiment. In particular, it is unclear why storage case 1 shows a much stronger effect than the more flexible cases 2 and 3.

As expected, we notice that there is hardly any reduction in the contract value if we compare the previous numerical experiment in Figure 3 with the current one (as long as we perform some protection on the volume boundary). This illustrates that both RBF2D and TBF2D based regressions are very stable methods, which function well even with very few simulations or with a very small number of scattered data sites. This leads us to conclude both methods in principle are very efficient multi-dimensional regression methods. We believe that this element will be key to the introduction of other dimensions to this valuation problem, especially on the volume level.

The second important point to note here is that both RBF2D and TBF2D based regression methods work indifferently on either grid or scattered data. This opens the opportunity to adaptively allocate the data sites where they are most needed, which is simply where the errors are the largest. This is only possible with methods working with scattered data. We believe this will prove beneficial again when we move to higher dimensional problems. The main lesson from this experiment is that during the process, one should maintain at least

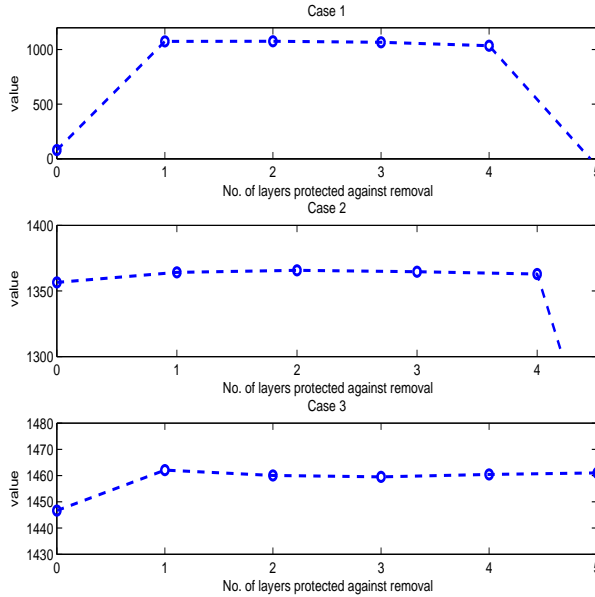


Figure 7: Impact of removing data sites using RBF2D regression

one layer of data sites on the volume level boundaries to aim for the maximum value for the contract. From a computational perspective, TBF2D benefits from working on a grid, whereas RBF2D does not.

6 Conclusion

In this article we introduced compactly supported Radial Basis Functions into the Least-Squares Monte Carlo regression setting. We employed two types of two-dimensional regression methods, one based on Radial Basis Functions in two dimensions (RBF2D) and the other one based on the Tensor of two univariate Basis Functions (TBF2D). We used these Least-Squares Monte Carlo methods to value gas storage with the spot approach. We proposed a variant of earlier spot approaches from Boogert & de Jong [1], and Carmona & Ludkovski [4] and Denault et al. [7] (CL/DSS). In our methodology we applied a two-dimensional regression on both price and volume, but maintained a uniform discretisation in the volume dimension whilst we refrained from using the mechanism introduced in CL/DSS to create ‘forward optimal paths’.

We compared the valuation of a slow (depleted field) and a fast (salt cavern) gas storage using our two-dimensional regression approaches against the one dimensional one from Boogert & de Jong [1] called Cubic1D. We found that both RBF2D and TBF2D provide a similar, yet consistently higher valuation than the Cubic1D for both the slow and the fast storage. All valuations are within

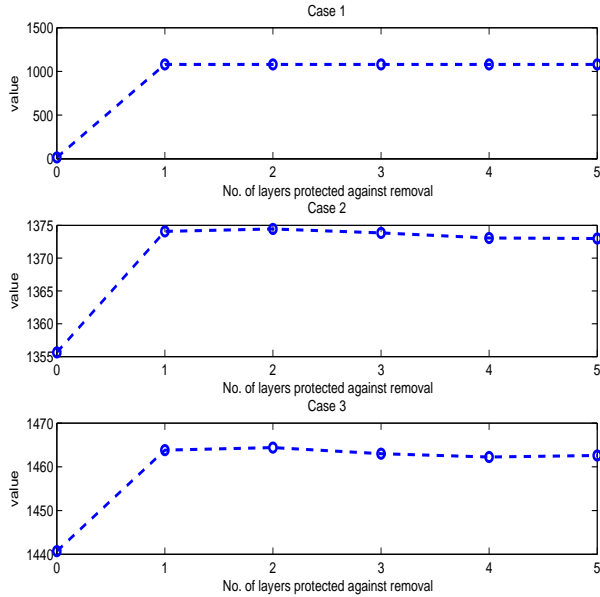


Figure 8: Impact of removing data sites using TBF2D regression

1% of each other. A positive point for the RBF2D and TBF2D based regression methods is that they both converge to the correct value for a lower number of simulations. If we compare RBF2D and TBF2D, we see TBF2D has several advantages. First, we find that TBF2D is computationally superior and a lot faster than the RBF2D due to the usage of the inversion method introduced by Hubbert & Mazières [12]. This indicates that TBF2D is more applicable to multiple dimensions. Another advantage concerns the shape factor. While RBF2D contains one shape factor which has to balance the support of two dimensions, the TBF2D contains one shape parameter for each dimension. Currently, we have employed a trial-and-error process to choose the shape factor(s). The natural next step is to investigate alternatives to choose the shape factor(s) in an automated fashion.

Next, we compared the impact of using scattered data in contrast to grid data. We found that if we take out price-volume data sites at random, the value can drop significantly. A solution to this problem is to protect the data sites on the volume boundaries. In our example, a single layer turns out to be sufficient. This finding draws our attention to the practice of randomisation employed in the CL/DSS approach in case the mechanism used to create forward optimal paths fails. Although one would expect that this randomisation method works, we suggest to change this complete randomisation to a procedure that ensures more data sites to be assigned at the volume level boundaries.

Compactly supported basis functions, which we used here, lead to a lower condition number for the regressor matrix compared to the one obtained from

globally supported functions. Besides, these basis offer us other benefits which we have not investigated yet. One of them is that they handle better discontinuities in the derivative of the value function. For example in a gas storage this is occurs when a penalty is applied in case the desired end volume is not reached.

In this article we limited ourselves to one price and one volume dimension. Previously it was shown that multi-factor price processes can be handled with a discretised approach (Boogert & De Jong [2]) and scattered approach (Carmona & Ludkovski [4]), while compactly supported RBF and TBF have been shown to work in multiple dimensions (Hubbert & Mazières [12]). As a next step, we plan to take up the problem of multiple price and volume dimensions. We believe that our new methodology in combination with the new regression methods will prove beneficial for that situation.

7 Acknowledgements

The authors like to thank Álvaro Cartea, Simon Hubbert, Henk Sjoerd Los, Simone Milani-Foglia, Zacharias Psaradakis, Oliver Salmon, Nicola Secomandi, Claire Valentine and the participants of the ESE Energy & Finance Conference 2011 for their input to the current paper.

References

- [1] Boogert, A. & C. De Jong, “Gas Storage Valuation using a Monte Carlo Method.” *Journal of Derivatives*, vol. 15 (3), 2008, p. 81-91.
- [2] Boogert, A. & C. De Jong, “Gas storage valuation using a multifactor price process.” *Forthcoming: Journal of Energy Markets*, vol 4 (4), 2011.
- [3] Buhmann, M.D., “Radial Basis Functions.” Cambridge Monographs on Applied and Computational Mathematics, 2003.
- [4] Carmona, R. & M. Ludkovski, “Valuation of energy storage: an optimal switching approach.” *Quantitative Finance*, vol. 10 (4), 2010, p. 359-374.
- [5] Carrière, J.F., “Valuation of the early-exercise price of options using simulations and nonparametric regression.” *Insurance: Mathematics and Economics*, vol. 19, 1996, p. 19–30.
- [6] Cheney, E.W. & W.A. Light, “Approximation Theory in Tensor Product Spaces.” Springer, 1985.
- [7] Denault, M., J.-G. Simonato & L. Stentoft, “A simulation-and-regression approach for stochastic dynamic programs with endogenous state variables.” Working paper G 2011 24, Les Cahiers du GERAD.

- [8] Fasshauer, G.F., "Meshfree Approximation Methods with Matlab." World Scientific, 2007.
- [9] Glasserman, P., "Monte Carlo Methods in Financial Engineering." Springer-Verlag, New York, 2004.
- [10] Grau, A.J., "Applications of Least-Squares Regressions to Pricing and Hedging of Financial Derivatives." PhD Dissertation, Technischen Universität München, 2008.
- [11] Horn, R.A. & C.R. Johnson, "Topics in matrix analysis." Cambridge University Press, 1991.
- [12] Hubbert, S. & D. Mazières, "Tensor product of compactly supported basis functions: Convergence order and reduction in computing time." Working paper, Birkbeck, University of London.
- [13] Longstaff, F.A. & E.S. Schwartz, "Valuing American Options by Simulation: A Simple Least-Squares Approach." *The Review of Financial Studies*, vol. 14 (1), 2001, p. 113-147.
- [14] Moreno, M. & J.F. Navas, "On the Robustness of Least-Squares Monte Carlo (LSM) for Pricing American Derivatives." *Review of Derivatives Research*, vol. 6 (2), 2003, p. 107-128.
- [15] Schaback, R., "Kernel-Based Meshless Methods", 2007, Available on: http://num.math.uni-goettingen.de/schaback/teaching/Appverf_II.pdf
- [16] Schwartz, E.S., "The stochastic behaviour of commodity prices: Implications for valuation and hedging." *Journal of Finance*, vol. 52 (3), 1997, p. 923-973.
- [17] Stentoft, L., "Assessing the Least Squares Monte Carlo approach to American option valuation." *Review of Derivatives Research*, vol. 7 (3), 2004, p. 129-168.
- [18] Tsitsiklis, J.N. & B. van Roy, "Regression Methods for Pricing Complex American-Style Options." *IEEE Transactions on Neural Networks*, vol. 12 (4), 2001, p. 694-703.
- [19] Wendland, H., "Scattered Data Approximation." Cambridge University Press, 2005.