Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 23/16

Anna Clevenhaus, Claudia Totzeck and Matthias Ehrhardt

(This is an updated/corrected version of Preprint 21/32)

A gradient based calibration method for the Heston model

October 11, 2023

http://www.imacm.uni-wuppertal.de

A gradient based calibration method for the Heston model

Anna Clevenhaus, Claudia Totzeck and Matthias Ehrhardt

2023-10-11

Abstract

The Heston model is a well-known two-dimensional financial model. Because the Heston model contains implicit parameters that cannot be determined directly from real market data, calibrating the parameters to real market data is challenging. In addition, some of the parameters in the model are non-linear, which makes it difficult to find the global minimum of the optimization problem within the calibration. In this paper, we present a first step towards a novel space mapping approach for parameter calibration of the Heston model. Since the space mapping approach requires an optimization algorithm, we focus on deriving a gradient descent algorithm. To this end, we determine the formal adjoint of the Heston PDE, which is then used to update the Heston parameters. Since the methods are similar, we consider a variation of constant and time-dependent parameter sets. Numerical results show that our calibration of the Heston PDE works well for the various challenges in the calibration process and meets the requirements for later incorporation into the space mapping approach. Since the model and the algorithm are well known, this work is formulated as a proof of concept.

1 Introduction

In finance, calibrating model parameters to fit real market data is challenging because most model parameters are implicit in the real market data [4, 7, 10, 12, 13, 15]. We consider the well-known two-dimensional Heston model, which contains at least four parameters implicit in the market data. The Heston proposed a two-dimensional stochastic differential equations (SDE) model to simulate the behavior of the stock price [5] and presented a closed-form valuation formula for his model. Some calibration techniques are based on this formula [13, 4]. The closed-form equation has some restrictions w.r.t. improving the model by considering non-constant parameters [13]. Another strategy to improve the model is to introduce additional processes, which leads to an increase in the difficulty of the calibration process [15]. Since the stock price and variance are stochastic processes in the Heston model, one can Monte Carlo optimization methods can be applied [15].

We consider the well-known two-dimensional Heston model, which contains at least four parameters implicit in the market data. The parameter calibration is formulated as

a constrained optimization problem to minimize a cost functional. The cost functional describes the difference between the reference data, the subsequent market data, and the data obtained by numerically solving our model. In the space mapping approach [1], a coarse and a fine optimization solver are used. This paper presents the first step "a gradient descent algorithm" for the Heston model towards a space mapping approach in finance. Later, within space mapping, the gradient descent algorithm will compute a coarse and a cheap approximation for the calibration problem of the Heston model.

Our goal is to introduce the space mapping approach to financial applications, and therefore we use the log-transformed normalized Heston model formulated as a partial differential equation (PDE) and the gradient descent algorithm as a pre-step, since both are well known and thus we can focus on the novel aspects. During our research, we didn't find any other PDE-based calibration approach for the Heston model. Of course, we are aware that there are already algorithms to compute the exact solution of the Heston calibration problem [4]. However, these methods are limited to the assumption that the parameters are constant, whereas our approach considers time-dependent parameters and constant parameter calibration. Furthermore, if the parameters are assumed to be time-dependent, an analytical solution can't be derived and thus an exact solution is not available. Therefore, as mentioned before, the paper is considered as a proof of concept to introduce a new calibration method based on financial research.

We focus on deriving a gradient descent algorithm for the Heston model that satisfies the requirements for its use in the space mapping approach. In addition to our financial model, the gradient descent algorithm requires the adjoint of the model. Therefore, we formally derive the adjoint of the Heston model and construct the gradient using well-known techniques from optimization of partial differential equations (PDEs) [6, 16]. The gradient descent algorithm has previously been applied to the Heston model using neural networks, so it has not been explicitly computed [11]. In this work, we focus on the calibration rather than the numerical solution of the model, i.e., our obtained results can be further improved by using more accurate numerical methods. Nevertheless, the calibration results obtained are remarkable. Our attempt can be easily extended to include time-dependent parameters for the variance, thus providing an improvement over current calibration strategies based on the semi-analytical solution of the Heston model.

The rest of the article is organized as follows. In Section 2, we introduce the Heston model and our approach to solving the resulting log-transformed Heston PDE. We then focus on parameter calibration. In Section 3, we derive the adjoint of the log-transformed Heston model and the corresponding gradient. We then discretize the method in Section 4 Finally, in Section 5 we focus on numerical results for the application of the gradient descent algorithm. We analyze the behavior of the algorithm in various problems for constant and time-dependent parameter calibration. The paper ends with a conclusion and an outlook in Section 6.

2 The Heston Model

The Heston model was proposed by Heston in 1993 [5] and describes the dynamics of the underlying asset $S \in \mathbb{R}^{>0}$ and the variance $\nu \in \mathbb{R}^{>0}$ by a two-dimensional SDE. This model is an extension of the well-known Black-Scholes model, which considers only a stochastic process for the asset and leads to

$$dS_t = rS_t\,dt + \sigma_S S_t\,dW_t^S, \quad S_0 > 0, \tag{1}$$

where $r \in \mathbb{R}^{>0}$ is the risk-free rate, $\sigma_S \in \mathbb{R}^{>0}$ is the volatility, and W_t^S is the Brownian motion. By definition, the variance is the square of the volatility of the asset, $\nu = \sigma_S^2$. Heston considered a Cox-Ingersoll process for modeling the variance, which leads to the SDE system of Heston's model under the risk-neutral measure given by

$$\begin{cases} dS_t = rS_t\,dt + \sqrt{\nu_t}S_t\,dW_t^S, & S_0 > 0, \\ d\nu_t = \kappa_\nu(\mu_\nu - \nu_t)\,dt + \sigma_\nu\sqrt{\nu_t}\,dW_t^\nu, & \nu_0 > 0, \end{cases}, \tag{2}$$

where $\kappa_\nu \in \mathbb{R}^{>0}$ is the mean reversion rate, $\mu_\nu \in \mathbb{R}^{>0}$ is the long term mean, and and $\sigma_\nu \in \mathbb{R}^{>0}$ is the volatility of variance. The Brownian motions W_t^S and W_t^ν are correlated by the constant $\rho \in [-1, 1]$ [5]. For the variance process to be positive, the Feller condition $2\kappa_\nu\mu_\nu \geq \sigma_\nu^2$ must be satisfied. If the Feller condition is violated, the (2) becomes complex which leads to computational problems. Using Kolmogorov's backward equation, we derive the Heston PDE under the risk-neutral measure

$$0 = V_t + \frac{\nu}{2}S^2V_{SS} + \frac{1}{2}\sigma_\nu^2\nu V_{\nu\nu} + rSV_S + \kappa_\nu(\mu_\nu - \nu)V_\nu + \sigma_\nu\nu\rho SV_{S\nu} - rV, \tag{3}$$

where $V(S, \nu, t)$ denotes the fair price of the option. If we look at the European plain vanilla put option, the terminal condition ("pay-off") is as follows

$$V(S, \nu, T) = (K - S, 0)^+, \tag{4}$$

where $K \in \mathbb{R}^{>0}$ denotes the predefined strike price. Heston proposed the following boundary conditions

$$V(0, \nu, t) = K\exp(-r(T - t)), \tag{5}$$
$$V(S, \nu, t) = 0, \quad \text{for } S \to \infty, \tag{6}$$
$$V(S, 0, t) = rsV_S(S, 0, t) + \kappa_\nu\mu_\nu V_\nu(S, 0, t) - rV(S, 0, t), \tag{7}$$
$$V(S, v, t) \sim K\exp(-r(T - t)), \quad \text{for } \nu \to \infty. \tag{8}$$

The boundary conditions for the underlying follow directly from the assumptions on the financial market. We reverse the time direction by introducing $\tau = T - t$ and thus the payoff condition (4) becomes an initial condition. Next, it is advantageous to use the variable transformation $x = \log(S/K)$ for the asset, as it simplifies the numerical implementation, i.e., we obtain a log-transformed variable with normalization to the strike

price K. We rewrite the option price equation as $\tilde{V}(x,\nu,\tau) = V(S,\nu,t)/K$ and obtain the so called log-*transformed normalized Heston PDE*

$$\tilde{V}_\tau = \frac{\nu}{2}\tilde{V}_{xx} + \frac{1}{2}\sigma_\nu^2\nu\tilde{V}_{\nu\nu} + (r - q - \frac{\nu}{2})\tilde{V}_x + \kappa_\nu(\mu_\nu - \nu)\tilde{V}_\nu + \sigma_\nu\nu\rho\tilde{V}_{x\nu} - r\tilde{V}, \tag{9}$$

defined on the semi-unbounded domain $x \in \mathbb{R}$, $\nu \geq 0$, $0 \leq \tau \leq T$ and supplied with the initial condition

$$\tilde{V}(x,\nu,\tau) = \Big(1 - \exp(x), 0\Big)^+ \tag{10}$$

and the following boundary conditions

$$\tilde{V}(x,\nu,\tau) \sim \exp(-r\tau), \quad \text{for } x \to -\infty, \tag{11}$$

$$\tilde{V}(x,\nu,\tau) = 0, \quad \text{for } x \to \infty, \tag{12}$$

$$\tilde{V}(x,\nu,\tau) \sim \exp(-r\tau), \quad \text{for } \nu \to \infty. \tag{13}$$

The boundary conditions for the underlying asset are the normalized log-transformation of Heston's proposed boundary conditions (5), (6) and (8). At $\nu = 0$ the parabolic PDE degenerates to a first order hyperbolic PDE, and therefore we need to consider the Fichera theory [2, 9] to assess whether it is necessary to provide these analytic boundary conditions or not.

From the Fichera condition at $\nu = 0$ given by

$$b(\nu) = \kappa_\nu(\mu_\nu - \nu) - \frac{1}{2}\sigma_\nu^2 \tag{14}$$

it follows

- if $\lim_{\nu\to 0^+} b(\nu) \geq 0$ (outflow) we must not supply any boundary conditions at $\nu = 0$.
- if $\lim_{\nu\to 0^+} b(\nu) < 0$ (inflow) we have to supply boundary conditions at $\nu = 0$.

In the sequel we assume that the Feller condition $2\kappa_\nu\mu_\nu \geq \sigma_\nu^2$ is satisfied. Therefore we obtain an outflow boundary at $\nu = 0$ and must not impose any analytical boundary conditions on this boundary. But we obviously need a numerical boundary condition to complete the scheme, which will be later discussed in section 4.

In addition to the Heston PDE with constant parameters for the variance process, we also consider the time-dependent parameters $\tilde{\kappa}_\nu$, $\tilde{\mu}_\nu$, $\tilde{\sigma}_\nu$, as well as $\tilde{\rho}$. Then the corresponding PDE formulation is given by

$$\tilde{V}_\tau = \frac{\nu}{2}\tilde{V}_{xx} + \frac{1}{2}\tilde{\sigma}_\nu^2\nu\tilde{V}_{\nu\nu} + (r - q - \frac{\nu}{2})\tilde{V}_x + \tilde{\kappa}_\nu(\tilde{\mu}_\nu - \nu)\tilde{V}_\nu + \tilde{\sigma}_\nu\nu\tilde{\rho}\tilde{V}_{x\nu} - r\tilde{V}. \tag{15}$$

In the numerical simulations 5, we will also discuss variations of constant and time-dependent parameter.

3 Derivation of a gradient based optimization strategy

For a given data set V_{data} and reference parameters ξ_{ref}, we formulate the calibration as optimization task with the cost functional given by

$$J(\tilde{V}, \xi) = \frac{1}{2} \int_0^T \|K\tilde{V} - V_{\text{data}}\|^2 \, d\tau + \frac{\lambda}{2} \|\xi - \xi_{\text{ref}}\|^2.$$

As our aim is to fit the parameter to real market data and therefore no ξ_{ref} exists, we set $\lambda = 0$ and the cost functional reduces to

$$J(\tilde{V}, u) = \frac{1}{2} \int_0^T \|K\tilde{V} - V_{\text{data}}\|^2 \, d\tau. \tag{16}$$

In the following, we derive a gradient-based algorithm that allows us to calibrate the parameters numerically. To this end we use a Lagrangian approach.

3.1 First-order optimality conditions for the Heston model

Let us denote the Lagrange multipliers by $\psi = (\varphi, \varphi^a, \varphi^b, \varphi^c, \varphi^d)$, set $\Omega = (0, \nu_{\max}) \times (x_{\min}, x_{\max})$ and split the boundary $\partial\Omega$ into

$$\begin{aligned} \Gamma_a &= \partial\Omega \cap \{x = -\infty\}, & \Gamma_b &= \partial\Omega \cap \{x = \infty\}, \\ \Gamma_c &= \partial\Omega \cap \{\nu = 0\}, & \Gamma_d &= \partial\Omega \cap \{\nu = \infty\}. \end{aligned}$$

First, we focus on the log-transformed normalized Heston equation with constant parameters (9) and define

$$A = \frac{1}{2}\nu \begin{pmatrix} \sigma_\nu^2 & \sigma_\nu \rho \\ \sigma_\nu \rho & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} \kappa_\nu(\mu_\nu - \nu) - \frac{1}{2}\sigma_\nu^2 \\ r - q - \frac{\nu}{2} - \frac{1}{2}\sigma_\nu \rho \end{pmatrix} \qquad \tilde{b} = \begin{pmatrix} \kappa_\nu \mu_\nu \\ r - q \end{pmatrix}.$$

Then it can be written as

$$\tilde{V}_\tau - \nabla \cdot A \nabla \tilde{V} - b \cdot \nabla \tilde{V} + r\tilde{V} = 0.$$

Next, we define the operator e which will represent the constraint in the Lagrangian. Since at Γ_c no boundary condition has to be given, we introduce $\tilde{\Omega} = \Omega \cap \Gamma$ and the operator e is implicitly defined by

$$\begin{aligned} \langle e(\tilde{V}, \xi), \psi \rangle = &\int_0^T \int_{\tilde{\Omega}} \Big[\tilde{V}_\tau - \nabla \cdot A \nabla \tilde{V} - b \cdot \nabla \tilde{V} + r\tilde{V} \Big] \varphi \, dz d\tau \\ &+ \int_0^T \int_{\Gamma_a} \Big[\tilde{V} - \exp(-r\tau) \Big] \varphi^a \, ds d\tau + \int_0^T \int_{\Gamma_b} \tilde{V} \varphi^b \, ds d\tau \\ &+ \int_0^T \int_{\Gamma_d} \Big[\tilde{V} - \exp(-r\tau) \Big] \varphi^d \, ds d\tau \\ =: &\, T_1 + T_2 + T_3 + T_4. \end{aligned} \tag{17}$$

The Lagrangian for the constrained parameter calibration problem is then given by

$$L(\tilde{V}, \xi, \psi) = J(\tilde{V}, \xi) - \langle e(\tilde{V}, \xi), \psi \rangle.$$

We formally compute the first-order optimality conditions by setting $dL = 0$. For details on the method we refer to [6, 16]. Before computing the Gâteaux derivatives of L in arbitrary directions [6], we note that by Green's first identity it holds

$$\int_\Omega (b \cdot \nabla \tilde{V})\varphi \, dz = \int_{\partial\Omega} (b \cdot \vec{n})\tilde{V}\varphi \, ds - \int_\Omega \tilde{V} \nabla \cdot (b\varphi) \, dz. \tag{18}$$

Therefore, we can rewrite

$$\begin{aligned} T_1 = &\int_0^T \int_{\tilde{\Omega}} \varphi \tilde{V}_\tau + A\nabla\tilde{V} \cdot \nabla\varphi - \frac{1}{2} b \cdot \nabla\tilde{V}\varphi + \frac{1}{2}\tilde{V} b \cdot \nabla\varphi + (r + \frac{1}{2}\nabla \cdot b)\tilde{V}\varphi \, dz \\ &- \int_{\partial\Omega} (A\nabla\tilde{V}) \cdot \vec{n}\varphi \, ds - \frac{1}{2}\int_{\partial\Omega} (b \cdot \vec{n})\tilde{V}\varphi \, ds d\tau \\ = &\Big[\int_\Omega \varphi\tilde{V} \, dz\Big]_{\tau=0}^{\tau=T} + \int_0^T \int_{\tilde{\Omega}} \tilde{V}\Big[-\varphi_\tau - \nabla \cdot A^\top \nabla\varphi + b \cdot \nabla\varphi + (r + \nabla \cdot b)\varphi\Big] \, dz \\ &+ \int_{\partial\Omega} [(A^\top \nabla\varphi) \cdot \vec{n} - (b \cdot \vec{n})\varphi]\tilde{V} \, ds - \int_{\partial\Omega} (A\nabla\tilde{V}) \cdot \vec{n}\varphi \, ds d\tau. \end{aligned}$$

As e is linear in V, the Gâteaux derivative in some arbitrary direction h reads

$$\begin{aligned} d_V \langle e(V, \xi), \psi \rangle[h] = &\Big[\int_\Omega \varphi h \, dz\Big]_{\tau=0}^{\tau=T} + \int_0^T \int_\Omega h\Big[-\varphi_\tau - \nabla \cdot A^\top \nabla\varphi + b \cdot \nabla\varphi + (r + \nabla \cdot b)\varphi\Big] dz \\ &+ \int_{\partial\Omega} \Big[(A^\top \nabla\varphi) \cdot \vec{n} - (b \cdot \vec{n})\varphi\Big] h \, ds - \int_{\partial\Omega} (A\nabla h) \cdot \vec{n}\varphi \, ds \, d\tau \\ &+ \int_0^T \int_{\Gamma_a} h\varphi^a \, ds \, d\tau + \int_0^T \int_{\Gamma_b} h\varphi^b \, ds \, d\tau + \int_0^T \int_{\Gamma_d} h\varphi^d \, ds \, d\tau. \end{aligned}$$

For the cost functional we have

$$d_{\tilde{V}} J(\tilde{V}, \xi)[h] = \int_0^T \int_\Omega hK(K\tilde{V} - V_{\text{data}}) \, dz \, d\tau.$$

To identify the adjoint equation, we consider

$$\begin{aligned} 0 &= d_{\tilde{V}} L(\tilde{V}, u, \psi)[h] \\ &= \int_0^T h \left[\int_\Omega K(K\tilde{V} - V_{\text{data}}) + \varphi_\tau + \nabla \cdot A^\top \nabla\varphi - b \cdot \nabla\varphi - (r + \nabla \cdot b)\varphi\right] dz \, d\tau. \end{aligned}$$

for arbitrary h. Note that we are not allowed to vary $\tilde{V}$ at $\tilde{V}(x, \nu, 0)$ as the initial condition is fixed. Therefore we have $h(0, \nu, x) \equiv 0$.

For choosing $h \equiv 0$ on $\partial\Omega$ and $h(T, \nu, x) = 0$, we find with the Variational Lemma

$$\varphi_\tau + \nabla \cdot A^\top \nabla\varphi - b \cdot \nabla\varphi - (r + \nabla \cdot b)\varphi = -K(K\tilde{V} - V_{\text{data}}) \quad \text{on } \Omega.$$

Now, choosing $h(T, \nu, x) \neq 0$, we then obtain the terminal condition $\varphi(T, \nu, x) = 0$.

We consider the four boundary conditions separately. At Γ_c, also the parabolic adjoint PDE degenerates to a first-order hyperbolic PDE, and thus we have to consider the Fichera theory [2, 9] for the variance again.

The Fichera condition w.r.t. the variance at $\nu = 0$ of the adjoint is the same as before. Therefore no analytic boundary condition is supplied for this boundary, as we assume that the Feller condition holds. On Γ_a we have

$$0 = \int_{\Gamma_a} \Big[(A^\top \nabla\varphi) \cdot \vec{n} - (b \cdot \vec{n})\varphi\Big] h - (A\nabla h) \cdot \vec{n}\varphi + h\varphi^a \, ds. \tag{19}$$

Choosing $h \equiv \text{const} \neq 0$ yields

$$0 = \int_{\Gamma_a} h\Big[(A^\top \nabla\varphi) \cdot \vec{n} - (b \cdot \vec{n})\varphi + \varphi^a\Big] \, ds,$$

hence $(A^\top \nabla\varphi) \cdot \vec{n} - (b \cdot \vec{n})\varphi + \varphi^a = 0$. On the other hand, choosing $\nabla h \neq 0$ (19) must still hold. This yields $\varphi = 0$ on Γ_a and

$$\varphi^a = -(A^\top \nabla\varphi) \cdot \vec{n} = -(A^\top \nabla\varphi) \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \frac{1}{2}\nu\sigma_\nu\rho\varphi_\nu + \frac{1}{2}\nu\varphi_x = \frac{1}{2}\nu\varphi_x$$

As $\tilde{V}(x_{\min}, \nu, 0) = \exp(-r\tau)$ is given and independent of x, we obtain $\varphi_x = 0$ there and thus $\varphi^a = \varphi = 0$ at this boundary. Similarly, we find on Γ_b that

$$0 = \int_{\Gamma_b} \Big[(A^\top \nabla\varphi) \cdot \vec{n} - (b \cdot \vec{n})\varphi\Big] h - (A\nabla h) \cdot \vec{n}\varphi + h\varphi_b \, ds. \tag{20}$$

With the same arguments, we obtain $\varphi = 0$ and $\varphi^b = 0$ on Γ_b as well. Following the same arguments for Γ_d we obtain $(A^\top \nabla\varphi) \cdot \vec{n} - (b \cdot \vec{n})\varphi + \varphi^d = 0$ with $\varphi = 0$ and thus it reduces to

$$\varphi^d = -(A^\top \nabla\varphi) \cdot \vec{n} = -(A^\top \nabla\varphi) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\frac{1}{2}\nu\sigma_\nu^2\varphi_\nu + \frac{1}{2}\nu\sigma_\nu\rho\varphi_x = -\frac{1}{2}\nu\sigma_\nu^2\varphi_\nu.$$

As $\tilde{V}(x, \infty, \tau) = \exp(-r\tau)$ is given and independent of ν, we obtain $\varphi_\nu = 0$ there and thus $\varphi^d = \varphi = 0$ at this boundary.

Altogether, the adjoint equation reads

$$\frac{\partial\varphi}{\partial\tau} + \nabla \cdot A^\top \nabla\varphi - b \cdot \nabla\varphi - (r + \nabla \cdot b)\varphi = -K(K\tilde{V} - V_{\text{data}}) \quad \text{on } \Omega, \tag{21}$$

which is equivalent to

$$\begin{aligned} \varphi_\tau + \frac{1}{2}\nu\sigma_\nu^2\varphi_{\nu\nu} + \nu\sigma_\nu\rho\varphi_{x\nu} + \frac{1}{2}\nu\varphi_{xx} + (\sigma_\nu^2 - \kappa_\nu(\mu_\nu - \nu))\varphi_\nu \\ + (q - r + \frac{\nu}{2} + \sigma_\nu\rho)\varphi_x + (\kappa_\nu - r)\varphi = -K(K\tilde{V} - V_{\text{data}}) \quad \text{on } \Omega \end{aligned} \tag{22}$$

with terminal condition $\varphi(T) = 0$ and $\varphi = 0$ on the boundaries Γ_a, Γ_b and Γ_d and the outflow boundary at $\nu = 0$.

3.2 Derivation of the Gradient

Let $\xi = (\sigma_\nu, \rho, \kappa_\nu, \mu_\nu)$ be the parameters to be identified, as r and q are given by the data. We compute the optimality condition by setting $d_\xi L(V, \xi, \psi) = 0$. Since the boundaries Γ_a, Γ_b and Γ_d are zero, we focus on $\tilde{\Omega}$. In the following we state the derivatives w.r.t. the different parameters separately. For σ_ν we get

$$d_{\sigma_\nu}\langle e(\tilde{V}, \xi), \psi\rangle = \int_0^T \int_{\tilde{\Omega}} \tilde{V}\Big[-\sigma_\nu \nu \varphi_{\nu\nu} - 2\sigma_\nu \varphi_\nu - \rho\varphi_x - \rho\nu\varphi_{x\nu}\Big]\, dz\, d\tau.$$

Similarly, we obtain for the other derivatives

$$\begin{aligned}
d_\rho\langle e(\tilde{V}, \xi), \psi\rangle &= \int_0^T \int_{\tilde{\Omega}} \tilde{V}\Big[-\sigma_\nu\varphi_x - \sigma_\nu\nu\varphi_{x\nu}\Big]\, dz\, d\tau,\\
d_{\kappa_\nu}\langle e(\tilde{V}, \xi), \psi\rangle &= \int_0^T \int_{\tilde{\Omega}} \tilde{V}\Big[(\mu_\nu - \nu)\varphi_\nu - \varphi\Big]\, dz\, d\tau,\\
d_{\mu_\nu}\langle e(\tilde{V}, \xi), \psi\rangle &= \int_0^T \int_{\tilde{\Omega}} \kappa_\nu \tilde{V}\varphi_\nu\, dz\, d\tau.
\end{aligned}$$

Note that $d_\xi L(\tilde{V}, \xi, \psi)[h_\xi] = 0$ needs to hold for arbitrary directions h_ξ. Therefore, we can read off the gradient from the above expressions.

We extend this gradient formulation for time dependent parameter $\tilde{\xi} = (\tilde{\sigma}_\nu, \tilde{\rho}, \tilde{\kappa}_\nu, \tilde{\mu}_\nu)$. The gradient is then time-dependent as well and given by

$$\begin{aligned}
d_{\sigma_\nu}\langle e(\tilde{V}, \tilde{\xi}), \psi\rangle &= \int_{\tilde{\Omega}} \tilde{V}\Big[-\tilde{\sigma}_\nu \nu \varphi_{\nu\nu} - 2\tilde{\sigma}_\nu \varphi_\nu - \tilde{\rho}\varphi_x - \tilde{\rho}\nu\varphi_{x\nu}\Big]\, dz\, d\tau\\
d_\rho\langle e(\tilde{V}, \tilde{\xi}), \psi\rangle &= \int_{\tilde{\Omega}} \tilde{V}\Big[-\tilde{\sigma}_\nu\varphi_x - \tilde{\sigma}_\nu\nu\varphi_{x\nu}\Big]\, dz\, d\tau,\\
d_{\kappa_\nu}\langle e(\tilde{V}, \tilde{\xi}), \psi\rangle &= \int_{\tilde{\Omega}} \tilde{V}\Big[(\tilde{\mu}_\nu - \nu)\varphi_\nu - \varphi\Big]\, dz\, d\tau,\\
d_{\mu_\nu}\langle e(\tilde{V}, \tilde{\xi}), \psi\rangle &= \int_{\tilde{\Omega}} \tilde{\kappa}_\nu \tilde{V}\varphi_\nu\, dz\,.
\end{aligned}$$

3.3 Gradient descent algorithm for the parameter calibration

Solving the first-order optimality condition all at once is difficult due to the forward-backward structure. Therefore, we propose a gradient descent algorithm in the following.

For a given initial parameter set ξ_0, we can solve the state equation for the Heston model with constant control variable ξ (9) or time dependent parameter (15). With the state solution at hand, we can compute the corresponding adjoint equation (21) or (22) backwards in time. Then we have all the information available to compute the gradient and update the parameter set using a gradient step. The procedure is sketched in algorithm 1.

Algorithm 1: Gradient descent method for Heston parameter calibration

Result: calibrated parameters for Heston model
initialize parameters u_0;
while $\|\text{gradient}\| > \epsilon$ **do**
  solve (9) or (15);
  solve (21) or (22);
  compute the gradient;
  line search for step size;
  update the parameter set;
end

Since the parameter domain for κ_ν, μ_ν, σ_ν and ρ is restricted, as well as the constraint that the Feller condition has to be fulfilled, we use the *projected Armijo rule* [16]. In the projected Armijo rule, we choose the maximum $\sigma_k \in \{1, 1/2, 1/4, \ldots\}$ for which

$$f\Big(\mathcal{P}(\xi_k - \sigma_k \nabla f(\xi_k))\Big) - f(\xi_k) \leq -\frac{\gamma}{\sigma_k}\|\mathcal{P}(\xi_k - \sigma_k \nabla f(\xi_k)) - \xi_k\|_2^2.$$

Here $\gamma \in (0, 1)$ is a numerical constant, which is problem-dependent and typically chosen as $\gamma = 10^{-4}$. We will use this value for the numerical results later on.

4 Discretization

In this section we introduce a closure boundary condition at $\nu = 0$ for the Heston and its adjoint and perform a domain truncation to discretize the problem. Following the Fichera theory and assuming that the Feller condition holds, no analytical boundary condition needs to be imposed at $\nu = 0$, neither for the Heston model nor for its adjoint, since the PDEs have a pure outflow boundary at this point. Nevertheless, we need a closure condition for this boundary for the discretization. We suggest to follow Heston's approach, as discussed in [9, 3] and use the reduced hyperbolic formulation of the Heston PDE and its adjoint. At Γ_c, we obtain

$$\tilde{V}_\tau = r\tilde{V}_x + \kappa_\nu \mu_\nu \tilde{V}_\nu - r\tilde{V} \tag{23}$$

for the log-transformed normalized PDE and similar for the adjoint

$$\varphi_\tau + (\sigma_\nu^2 - \kappa_\nu \mu_\nu)\varphi_\nu + (q - r + \sigma_\nu \rho)\varphi_x + (\kappa_\nu - r)\varphi = -K(K\tilde{V} - V_{\text{data}}). \tag{24}$$

Now, we perform a domain truncation to obtain a rectangular grid, instead of a semi-unbounded domain and introduce the grid points. We consider uniform meshes in each direction and obtain for the spatial directions $x_i = x_{\min} + i\Delta_x$ for $i = 0, \ldots, N_x$ with $\Delta_x = (x_{\max} - x_{\min})/N_x$ and $\nu_j = j\Delta_\nu$ for $j = 0, \ldots, N_\nu$ with $\Delta_\nu = \nu_{\max}/N_\nu$, as well as $\tau_k = k\,\Delta_\tau$ for $k = 0, \ldots, N_\tau$ with $\Delta_\tau = T/N_\tau$ for the temporal direction.

Since this is a proof-of-concept, simple and well-known spatial and temporal discretization methods are used to illustrate our approach. For the time discretization, we use the

well-known *alternating direction implicit* (ADI) method, in more detail the Hundsdorfer-Verwer scheme [8]. This scheme is of second order for any choice of θ, where θ is a measure of classification similar to the θ-method, and is able to handle mixed derivative terms. To present the ADI method, we use a general second order PDE formulation

$$u_\tau + a_{11}u_{\nu\nu} + 2a_{12}u_{x\nu} + a_{22}u_{xx} + b_1u_\nu + b_2u_x + cu = 0. \tag{25}$$

Since the log-transformed normalised Heston PDE and its adjoint are second-order PDEs and we further use the same discretization.

In a first step, we split the operator of the PDE into three operators

$$\mathcal{F}(\tau) = \mathcal{F}_0(\tau) + \mathcal{F}_1(\tau) + \mathcal{F}_2(\tau), \tag{26}$$

with

$$\begin{aligned}
\mathcal{F}_0(\tau) &= 2a_{12}D_{x\nu} \\
\mathcal{F}_1(\tau) &= b_2D_x + a_{22}D_{xx} - \frac{1}{2}cI \\
\mathcal{F}_2(\tau) &= b_1D_\nu + a_{11}D_{\nu\nu} - \frac{1}{2}cI,
\end{aligned}$$

where D_x describes the discretization matrix of the first derivative w.r.t. x and accordingly D_{xx} of the second derivative w.r.t. x, D_ν and $D_{\nu\nu}$ of the first and second derivative w.r.t. ν, I denotes the identity matrix. The discretization matrices are derived using central finite differences. Let $u_k^{i,j} \approx u(x_i, \nu_j, \tau_k)$ and simplify $u_k \approx u(x, \nu, \tau_k)$. In each time step, we have to solve the following system of equations

$$\begin{cases}
Y_0 = u_k + \Delta\tau\mathcal{F}(\tau_k)u_k, \\
Y_1 = Y_0 + \theta\Delta\tau\Big(\mathcal{F}_1(\tau_{k+1})Y_1 - \mathcal{F}_1(\tau_k)u_k\Big), \\
Y_2 = Y_x + \theta\Delta\tau\Big(\mathcal{F}_2(\tau_{k+1}, Y_2) - \mathcal{F}_2(\tau_k)u_k\Big), \\
\tilde{Y}_0 = Y_0 + \frac{1}{2}\Delta\tau\Big(\mathcal{F}(\tau_{k+1})Y_2 - \mathcal{F}(\tau_k)u_k\Big), \\
\tilde{Y}_1 = \tilde{Y}_0 + \theta\Delta\tau\Big(\mathcal{F}_1(\tau_{k+1})\tilde{Y}_1 - \mathcal{F}_1(\tau_k)u_k\Big), \\
\tilde{Y}_2 = \tilde{Y}_x + \theta\Delta\tau\Big(\mathcal{F}_2(\tau_{k+1})\tilde{Y}_2 - \mathcal{F}_2(\tau_k)u_k\Big), \\
u_{k+1} = \tilde{Y}_2.
\end{cases} \tag{27}$$

We choose $\theta = 3/4$ and improve the implementation as we used the approach in [14] by using a matrix based instead of a vector based implementation of u_k.

At this point, we have to discuss the boundary conditions for the Heston model again. We start with the x dimension. We set $\tilde{V}(x_{\min}, \nu, \tau) = \exp(-r\tau)$ and similar $\tilde{V}(x_{\max}, \nu, \tau) = 0$, as we suggest a sufficiently small $x_{\min}$ and a sufficiently large $x_{\max}$. For the variance boundaries, we follow the approach of Kùtik and Mikula [9]. Due to the Fichera theory, at $\nu = 0$ we gain a outflow boundary and no information can enter the domain from the region $\nu < 0$. Since the same holds for the adjoint, we propose the same approach there. In these, we extend the numerical domain for this case and introduce

ghost cells $p_{i,0} = (x_i, \nu_0)$, $i = 1, \dots, N$ where $\nu_0 = -\Delta_y$ and determine the ghost cell values $u^n_{p_{i,0}}$, $i = 1, \dots, N$ by zero order extrapolation. We obtain a constant function

$$u^n_{p_{i,0}} = u^n_{p_{i,1}}, \quad i = 1, \dots, N. \tag{28}$$

Now we consider the boundary, where $\nu \to \infty$ and the truncation shrinks the domain from $(0, \infty) \to (0, \nu_{\max})$. Dirichlet boundary condition from Heston would causes an unnatural jump in the solution. Therefore different strategies have been developed to overcome this issue, e.g. [9]. Again, we follow Kùtik and Mikula and impose artificial homogeneous Neumann boundary conditions. The choice is motivated by the variance independence of the original boundary condition from Heston, for sufficiently large $\nu_{\max} = \mathcal{O}(1)$. Again, we use zero order extrapolation to implement this condition. We add ghost cells $p_{i,N_\nu+1}$, $i = 1, \dots, N$ where $\nu_{N_\nu+1} = \nu_{\max} + \Delta_y$ and determine their values $u^n_{p_{i,N_\nu+1}}$, $i = 1, \dots, N$ by a constant function

$$u^n_{p_{i,N_\nu+1}} = u^n_{p_{i,N_\nu}}, \quad i = 1, \dots, N. \tag{29}$$

For a discussion about different numerical boundary conditions for the variance of the Heston model, we refer to [3]. The integrals appearing in the gradient are computed by the trapezoidal rule.

5 Numerical Results

Following [3], we use

$$K = 1.0, r = 0.1, \xi_{\text{ref}} = (5.0, 0.07, 0.5, -0.5) \tag{30}$$

for generating an artificial V_{data} for each time step τ_k. For the discretization, we use $N_x = 79, N_\nu = 39, N_\tau = 59$. As bounds for the projected Armijo-rule for $\xi = (\kappa_\nu, \mu_\nu, \sigma_\nu, \rho)$, we set $0 < \kappa_\nu < 8, \quad 0 < \mu_\nu < 1, \quad 0 < \sigma_\nu < 1, \quad -1 < \rho < 1$. Note that the projected Armijo-rule ensures that the Feller condition holds within each optimization step. Hence we are in the case of an outflow boundary. We set the maximal iteration value for the calibration to 20. For the initial guesses $\xi_{\text{init}} = (\kappa^{\text{init}}_\nu, \mu^{\text{init}}_\nu, \sigma^{\text{init}}_\nu, \rho^{\text{init}})$ we used generated random numbers within a maximal percentage difference from ξ_{ref}. We use four different percentages $10, 25, 50$ and 75 and for each we generate five sets. The initial parameters as well as the calibrated parameters in the constant case are given in Table 1 for κ_ν, Table 2 for μ_ν, Table 3 for σ_ν and Table 4 for ρ.

We observe that the calibration leaves κ_ν nearly untouched, whereas the other parameter values are significantly changed. This could be reasoned by the structure of the drift term $\kappa_\nu(\mu_\nu - \nu)$, since κ_ν and μ_ν are multiplied. As a optimization measure, we compute the relative reduction of the cost functional using

$$r(\xi_{\text{init}}) = 100 \cdot \left(1 - \frac{J(\tilde{V}, \xi_{\text{opt}})}{J(\tilde{V}, \xi_{\text{init}})}\right). \tag{31}$$

Table 5 shows the relative reduction of the cost functional for the different test cases for ξ_{init}.

κ_ν^{init}	T1	T2	T3	T4	T5	κ_ν^{cal}	T1	T2	T3	T4	T5
10	4.60	5.15	4.50	4.70	4.50	10	4.60	5.15	4.51	4.70	4.50
25	4.75	4.60	5.25	5.90	4.45	25	4.75	4.60	5.24	5.88	4.45
50	4.10	3.25	3.55	2.90	7.00	50	4.11	3.26	3.55	2.94	6.95
75	5.70	7.30	7.00	7.50	6.60	75	5.69	7.29	6.85	7.40	6.60

Table 1: Five initial values sets for κ_ν^{init} for the different percentages and the corresponding calibrated values for C0.

μ_ν^{init}	T1	T2	T3	T4	T5	μ_ν^{cal}	T1	T2	T3	T4	T5
10	0.0679	0.0637	0.0735	0.0714	0.0714	10	0.0598	0.0727	0.0575	0.0632	0.0577
25	0.0721	0.0777	0.0616	0.0630	0.0868	25	0.0655	0.0610	0.0720	0.0836	0.0615
50	0.1001	0.0406	0.0840	0.0448	0.0434	50	0.0476	0.0148	0.0484	0.0136	0.0882
75	0.0742	0.0532	0.0448	0.0399	0.0490	75	0.0847	0.1069	0.1005	0.1064	0.0928

Table 2: Five initial values sets for μ_ν^{init} for the different percentages and the corresponding calibrated values for C0.

σ_ν^{init}	T1	T2	T3	T4	T5	σ_ν^{cal}	T1	T2	T3	T4	T5
10	0.465	0.545	0.510	0.515	0.480	10	0.451	0.559	0.484	0.502	0.455
25	0.575	0.570	0.505	0.415	0.480	25	0.565	0.546	0.523	0.455	0.429
50	0.580	0.360	0.630	0.345	0.655	50	0.501	0.292	0.581	0.261	0.713
75	0.705	0.470	0.235	0.350	0.565	75	0.718	0.562	0.390	0.489	0.630

Table 3: Five initial values sets for σ_ν^{init} for the different percentages and the corresponding calibrated values for C0.

ρ^{init}	T1	T2	T3	T4	T5	ρ^{cal}	T1	T2	T3	T4	T5
10	-0.550	-0.485	-0.520	-0.520	-0.495	10	-0.483	-0.547	-0.404	-0.460	-0.384
25	-0.410	-0.490	-0.605	-0.585	-0.380	25	-0.370	-0.390	-0.690	-0.829	-0.159
50	-0.555	-0.465	-0.350	-0.685	-0.725	50	-0.258	-0.242	-0.207	-0.441	-0.847
75	-0.695	-0.205	-0.775	-0.655	-0.150	75	-0.743	-0.972	-0.990	-0.990	-0.437

Table 4: Five initial values sets for ρ^{init} for the different percentages and the corresponding calibrated values for C0.

C0	T1	T2	T3	T4	T5
10	43.12	77.38	65.22	63.22	52.75
25	58.46	78.38	62.14	66.06	64.53
50	83.09	20.84	52.76	30.98	72.62
75	16.23	72.89	82.89	82.09	80.43

Table 5: Relative reduction of the cost functional computed with 31 using the different test cases for ξ_{init} and ξ_{cal} in the constant calibration setting.

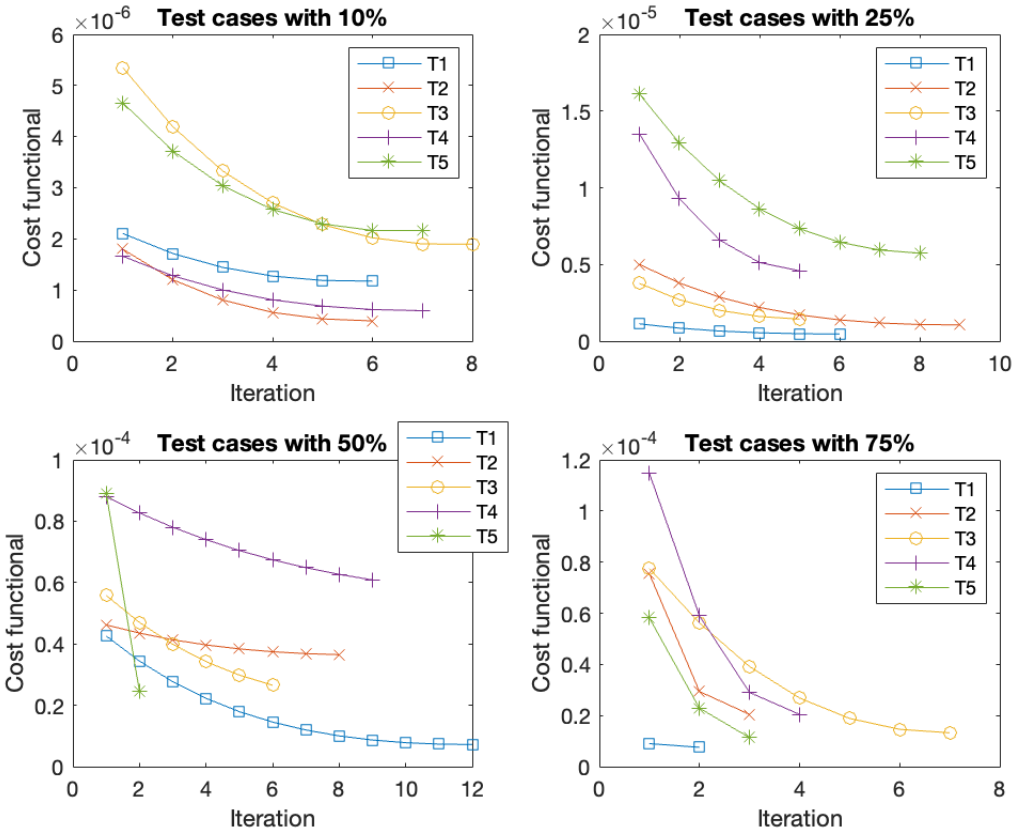

Figure 1: Cost functional evolution per iteration for the test cases within the constant parameter calibration.

	κ_ν	μ_ν	σ_ν	ρ
C0	constant	constant	constant	constant
C1	time-dependent	constant	constant	constant
C2	constant	time-dependent	constant	constant
C3	constant	constant	time-dependent	constant
C4	constant	constant	constant	time-dependent
C5	time-dependent	time-dependent	time-dependent	time-dependent
C6	constant	time-dependent	time-dependent	time-dependent
C7	constant	time-dependent	constant	time-dependent

Table 6: Different cases for calibration setting.

C1	T1	T2	T3	T4	T5
10	44.48	78.75	64.91	64.12	54.03
25	58.97	78.98	62.96	66.77	63.96
50	83.13	19.59	45.71	33.45	72.84
75	16.59	72.80	82.14	81.13	80.94

Table 7: Relative cost function reduction for the C1 calibration.

As the calibrated values differ over the test cases, it is reasonable to assume, that we only find local minima. However the results are remarkable.

Since in the real market the parameter are not considered as constant, we improve the approach by considering different parameters, and some parameter sets as time-dependent. From the relative change in the constant calibration setting, we choose the following (additional) test cases listed in Table 6. Further the table includes the links to the cost functional reduction tables and figures of the cost function evolution with respect to the different cases.

For each case of the time-dependent test cases V_{data} is generated as before and ξ^{init} is assumed to be constant and thus also the same as before.

To quantify the different calibrations, we compute the average relative cost functional reduction by

$$r_a(\xi_{\text{init}}) = \frac{r(\xi_{\text{init}})}{20} \tag{32}$$

and summarize the results in Table 14. Note, that all cost function reduction averages are huge. We observe that a time-dependent calibration for one parameter alone (C1-C4)

C2	T1	T2	T3	T4	T5
10	42.96	77.85	63.17	62.09	53.05
25	58.03	77.05	63.09	68.46	70.00
50	82.90	20.79	51.89	29.79	73.66
75	16.05	74.64	82.95	82.43	78.89

Table 8: Relative cost function reduction for the C2 calibration.

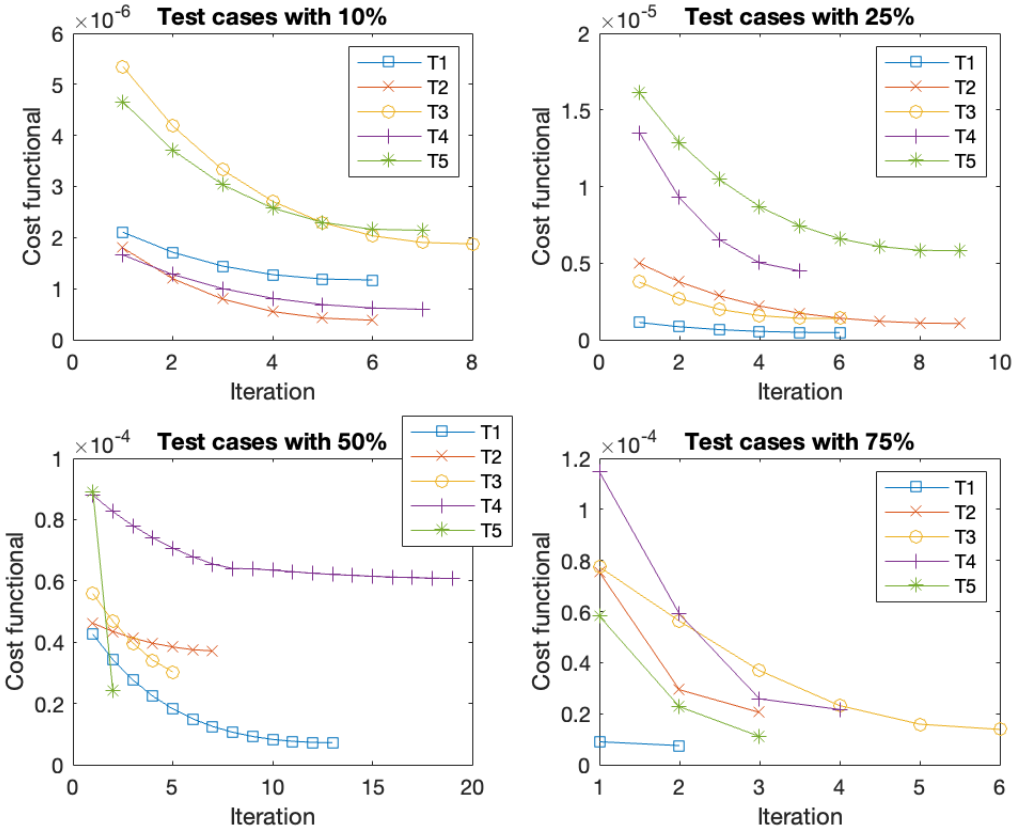

Figure 2: Cost functional evolution per iteration for the different test set in the case scenario C1.

C3	T1	T2	T3	T4	T5
10	44.22	77.89	64.50	64.04	53.52
25	58.49	78.76	61.62	65.92	64.50
50	83.05	20.97	52.49	30.55	72.23
75	15.96	72.91	82.71	82.10	80.44

Table 9: Relative cost function reduction for the C3 calibration.

C4	T1	T2	T3	T4	T5
10	49.03	75.90	66.62	66.49	56.81
25	60.42	77.71	62.64	68.50	62.96
50	78.54	18.24	52.63	30.29	74.03
75	20.89	73.59	81.75	80.49	79.74

Table 10: Relative cost function reduction for the C4 calibration.

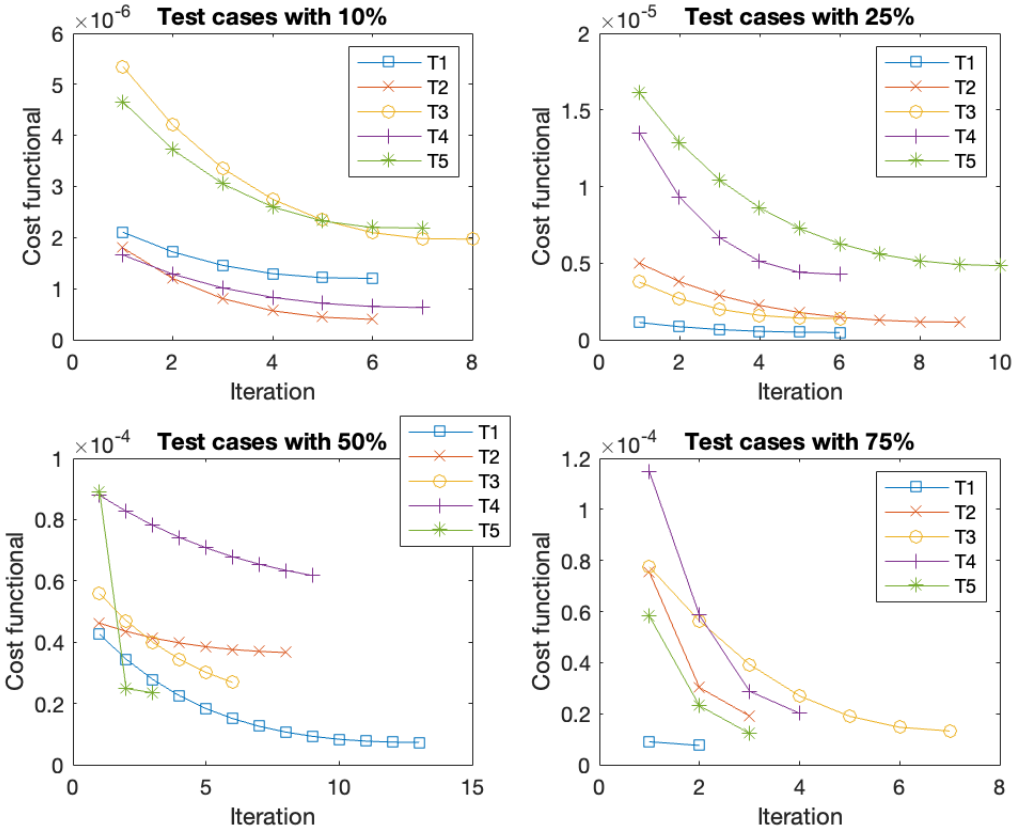

Figure 3: Cost functional evolution per iteration for the different test set in the case scenario C2.

C5	T1	T2	T3	T4	T5
10	47.63	75.50	65.79	64.45	57.08
25	60.70	76.67	62.36	68.51	71.33
50	82.32	20.89	58.63	31.72	72.96
75	20.79	73.10	78.14	77.71	78.76

Table 11: Relative cost function reduction for the C5 calibration.

C6	T1	T2	T3	T4	T5
10	47.91	76.08	65.50	64.77	56.43
25	60.27	76.39	63.37	70.16	68.91
50	80.94	19.19	57.15	29.38	73.71
75	20.84	74.74	81.70	80.74	79.10

Table 12: Relative cost function reduction for the C6 calibration.

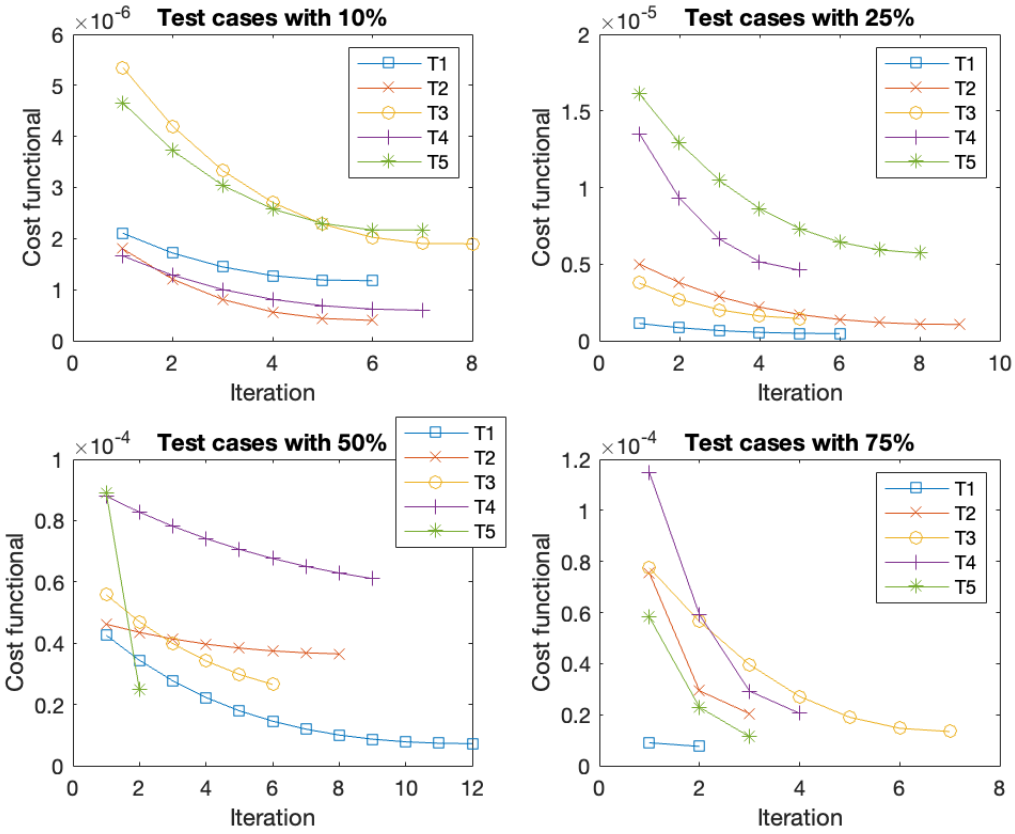

Figure 4: Cost functional evolution per iteration for the different test set in the case scenario C3.

C7	T1	T2	T3	T4	T5
10	47.79	75.66	65.50	64.66	56.48
25	60.21	76.37	62.80	69.43	69.29
50	81.08	19.35	57.16	30.05	73.02
75	20.62	74.55	80.59	80.09	78.89

Table 13: Relative cost function reduction for the C7 calibration.

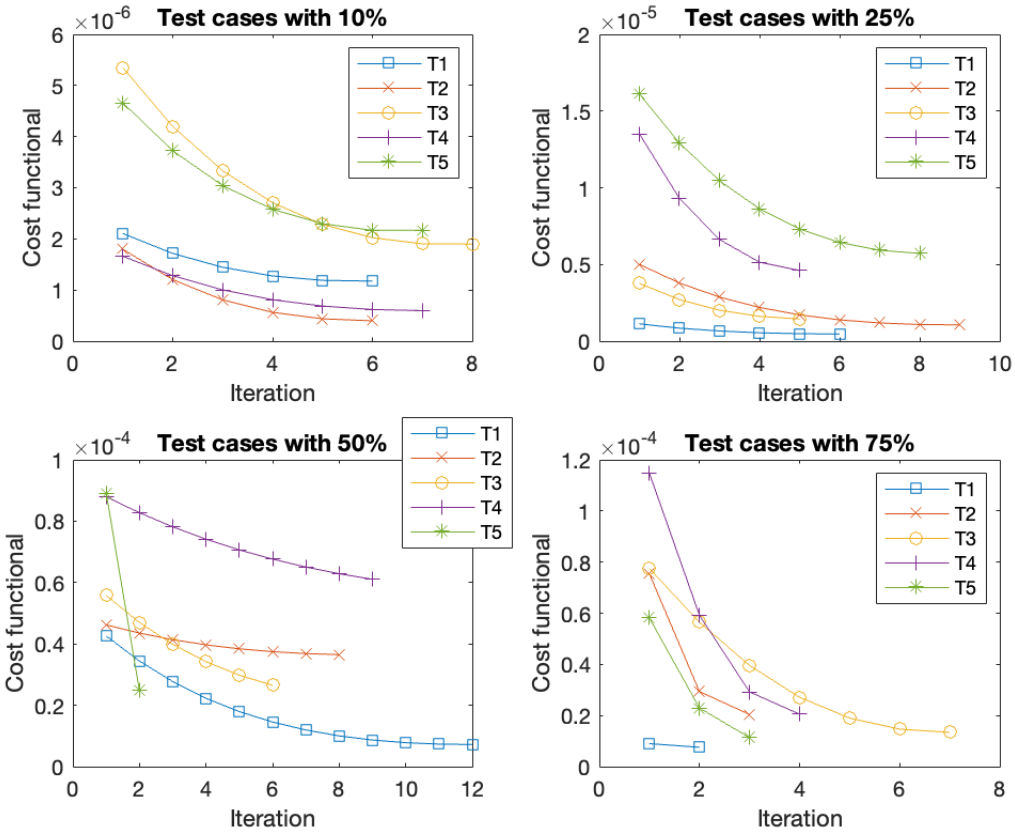

Figure 5: Cost functional evolution per iteration for the different test set in the case scenario C4.

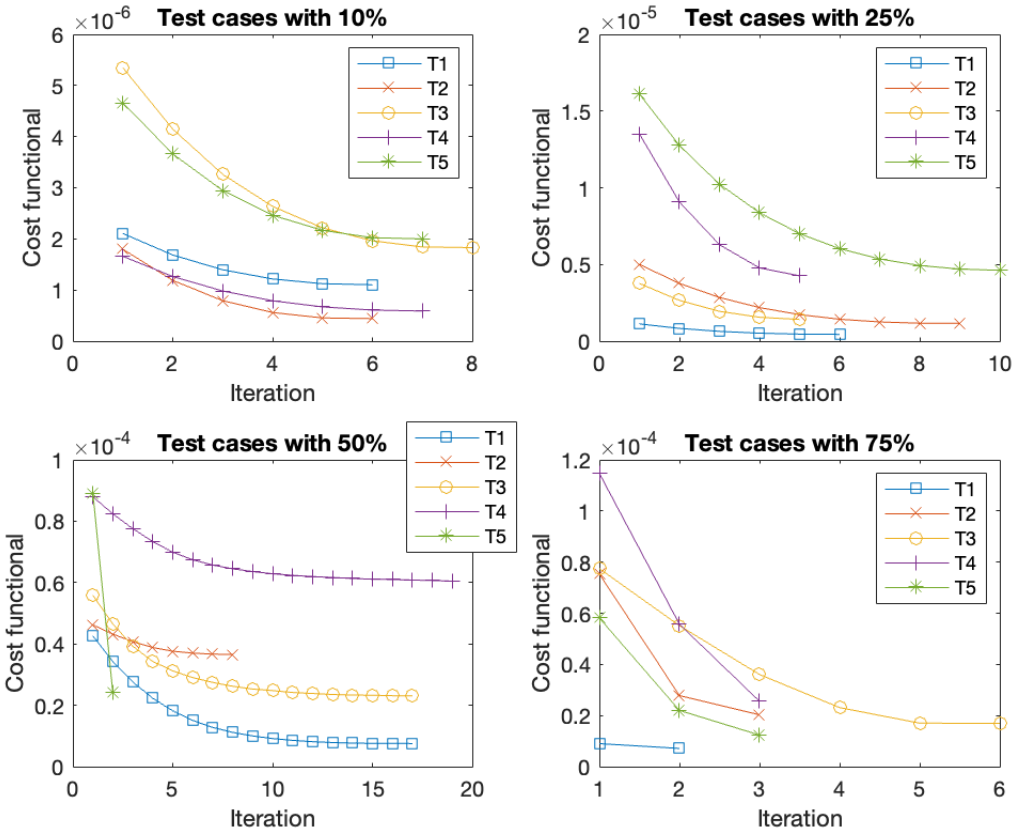

Figure 6: Cost functional evolution per iteration for the different test set in the case scenario C5.

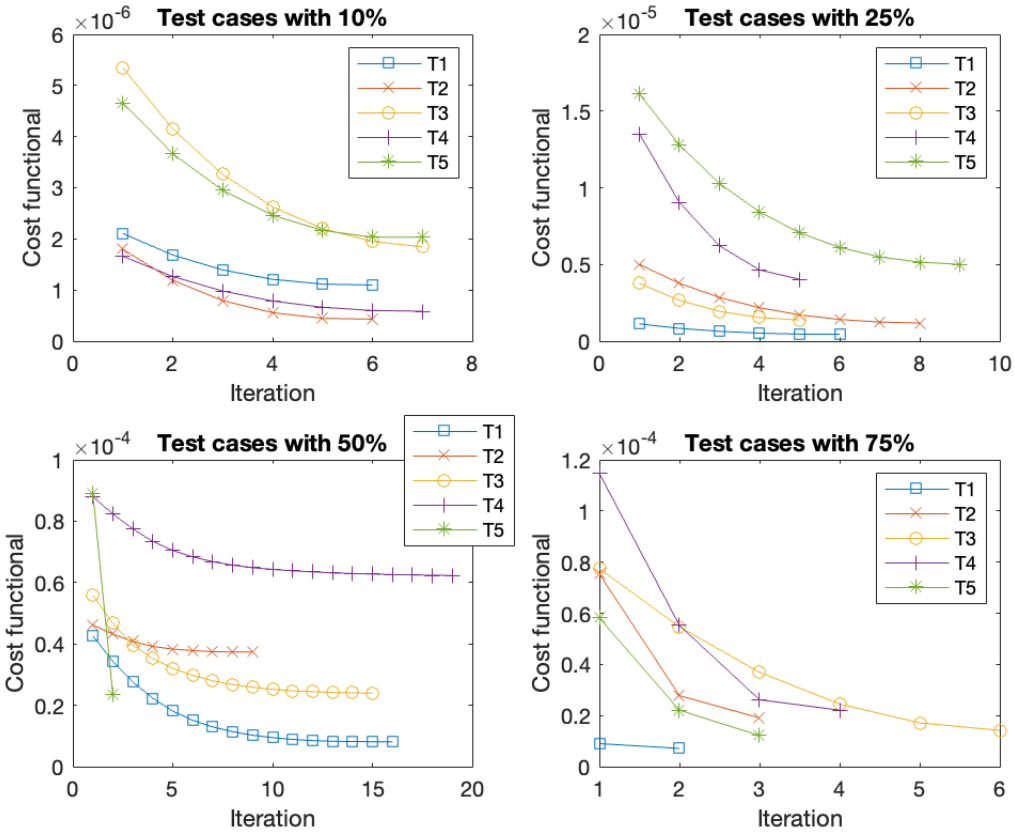

Figure 7: Cost functional evolution per iteration for the different test set in the case scenario C6.

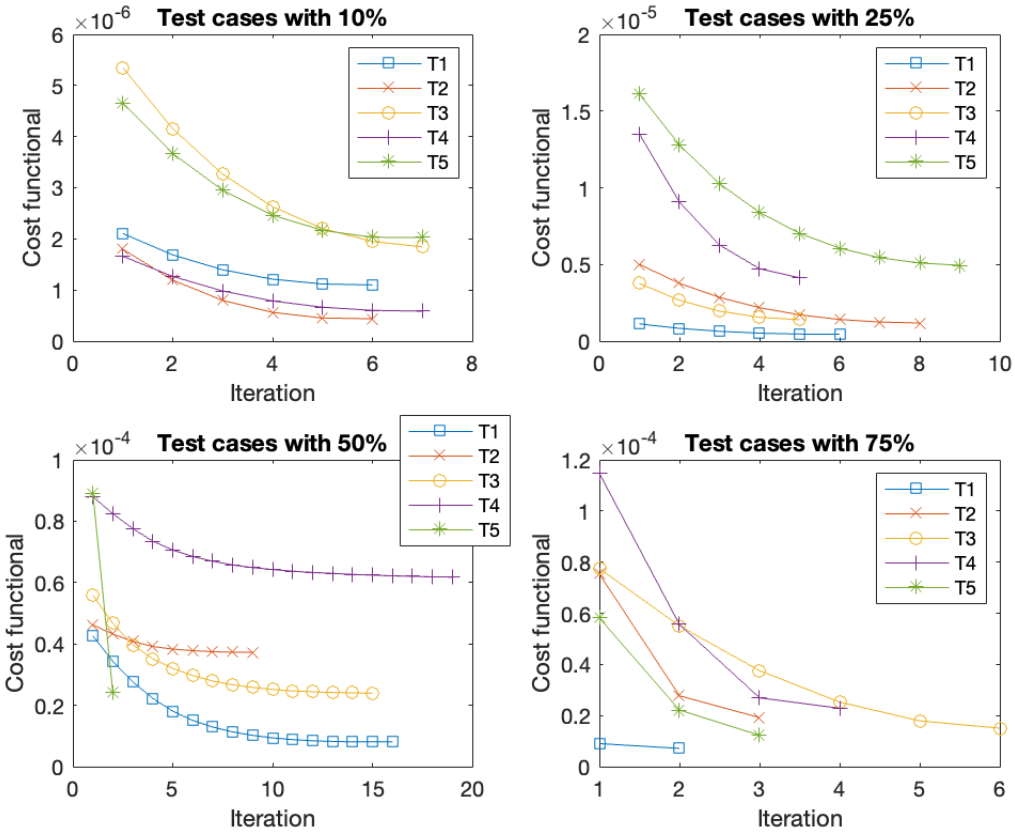

Figure 8: Cost functional evolution per iteration for the different test set in the case scenario C7.

Case Setup	C0	C1	C2	C3	C4	C5	C6	C7
$r_a(\xi_{\text{init}})$	61.30	61.31	61.49	61.34	61.86	62.25	62.36	66.12
table	5	7	8	9	10	11	12	13
figure	1	2	3	4	5	6	7	8

Table 14: Average cost functional reduction in percentage for the different cases over the 20 test cases, as well as a list for the corresponding relative cost functional reduction as well as the figures of the cost functional evolution.

doesn't improve the cost functional reduction significantly. The first slightly improvement can be found by using at least two time-dependent parameters (C5-C7). Surprisingly C7 gives the best calibration results, where κ_ν is the only variable which is calibrated as a constant, and C5, where all parameters are calibrated as time-dependent gives the least improvement, when considering combinations of time-dependent parameter calibration. The fact that V_{data} is generated with constant parameters and the best case considers time-dependent parameters indicates that the time-dependency is a good way to overcome the local minima. Those results are inline with literature [13, 15].

6 Conclusion and Outlook

Our paper began with an introduction to the Heston model. The introduction was followed by the derivation of a gradient-based optimization, including the derivation of a gradient descent algorithm. In the next section, the discretization of the schemes and algorithms is presented and numerical results are illustrated. These show that the gradient descent algorithm is a feasible calibration technique. The relative cost function reduction is remarkable, even though we can expect to find only local minima. Thus, our approach follows recent research. Furthermore, the assumption of at least time-dependent parameters is better suited to the real market situation, since in the real market almost no parameter is constant. In future work, we plan to incorporate the gradient descent algorithm into the space mapping approach and to test the space mapping approach on real market data.

References

[1] J.W. Bandler, Q.S. Cheng, S.A. Dakroury, A.S. Mohamed, M.H. Bakr, K. Madsen, and J. Sondergaard. Space mapping: the state of the art. *IEEE Trans. Microw. Theory Techn.*, 52(1):337–361, 2004.

[2] Z. Bučková, M. Ehrhardt, and M. Günther. Fichera theory and its application in finance. In G. Russo, V. Capasso, G. Nicosia, and V. Romano, editors, *Progress in Industrial Mathematics at ECMI 2014*, pages 103–111, Cham, 2016. Springer International Publishing.

[3] A. Clevenhaus, C. Totzeck, and M. Ehrhardt. A study of the effects of different boundary conditions on the variance of the Heston model. In K. Burnecki and M. Teuerle, editors, *Progress in Industrial Mathematics at ECMI 2023*. Springer International Publishing, 2024.

[4] Y. Cui, S. del Baño Rollin, and G. Germano. Full and fast calibration of the Heston stochastic volatility model. *Eur. J. Oper. Res.*, 263 (2):625–638, 2017.

[5] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Fin. Stud.*, 6(2):327–343, 1993.

[6] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints*. Springer, 2009.

[7] B. Horvath, A. Muguruza, and M. Tomas. Deep learning volatility: a deep neural network perspective on pricing and calibration in (rough) volatility models. *Quantitative Finance*, 21(1):11–27, 2021.

[8] W. Hundsdorfer. Accuracy and stability of splitting with stabilizing corrections. *Appl. Numer. Math.*, 42(1-3):213–233, 2002.

[9] P. Kútik and K. Mikula. Diamond–cell finite volume scheme for the Heston model. *(DCDS-S*, 8(5):913–931, 2015.

[10] S. Liu, A. Borovykh, L.A. Grzelak, and C.W. Oosterlee. A neural network-based framework for financial model calibration. *J. Math. Ind.*, 9:1–28, 2019.

[11] S. Liu, C.W. Oosterlee, and S. M. Bohte. Pricing options and computing implied volatilities using neural networks. *Risks*, 7(1):16, 2019.

[12] F. Mehrdoust, I. Noorani, and A. Hamdi. Two-factor Heston model equipped with regime-switching: American option pricing and model calibration by Levenberg-Marquardt optimization algorithm. *Math. Comput. Simul.*, 204:660–678, 2023.

[13] S.S. Mikhailov and U. Nögel. Heston's stochastic volatility model implementation, calibration and some extensions. John Wiley and Sons, 2004.

[14] L. Teng and A. Clevenhaus. Accelerated implementation of the ADI schemes for the Heston model with stochastic correlation. *J. Comput. Sci.*, 36:101022, 2019.

[15] L. Teng, M. Ehrhardt, and M. Günther. On the Heston model with stochastic correlation. *Int. J. Theor. Appl. Fin.*, 6, 2016.

[16] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Springer, 2009.