Gaussian Processes for Implied Volatility Estimation

Tejas Appana, Andrew Carranti, Arjun Koshal, Adam Moszczynski, Matthew Thomas

April 26, 2024

Contents

1 Abstract

The accurate modeling of the implied volatility surface holds great importance in option pricing, trading, and hedging. Significant research has been devoted to this task, ranging from mathematical models to machine-learning solutions. However, these approaches generate point estimates based on imperfect market data, which can lead to inaccurate and misleading models. Gaussian Processes, a Bayesian non-parametric method, have addressed this challenge. They generate posterior distributions of implied volatility functions, producing confidence intervals and quantifiable measures of model uncertainty. Our model is applied to end-of-day option quotes, showing high accuracy that can outperform traditional parametric models in implied volatility estimation and option pricing. Our findings could significantly impact hedging strategies, as the distribution that our model derives accounts for a spectrum of scenarios, enabling a more nuanced assessment of risk. We expect this to demonstrate the effectiveness of our approach in reducing risk exposure compared to traditional methods, thereby providing a tangible measure of the practical benefits of incorporating our Gaussian Process method in the real world.

2 Introduction

The dynamic nature of financial markets has made efficiently modeling and predicting asset prices a fundamental challenge in the domains of derivatives pricing, risk management, and finance as a whole. At the heart of derivatives valuation lies the implied volatility surface, a fundamental construct that describes market expectations of future volatility across various strike prices and maturities. Accurately modeling and calibrating this surface is critical for pricing options and constructing robust hedging strategies. Traditional parametric models often struggle to capture the complex behaviors exhibited by market movements, such as stochastic volatility and price jumps. However, purely data-driven approaches, while flexible, often lack the robustness required for precise, out-of-sample predictions.

Gaussian processes (GPs) have recently gained attention as a powerful alternative for this task, offering a versatile, Bayesian non-parametric framework for approximating the complex functions that describe market expectations in implied volatility surfaces. This approach not only allows for more accurate predictions of volatility but also quantifies uncertainty in these predictions, an attribute critical for risk management. The inherent flexibility of Gaussian processes, free from rigid functional form assumptions, allows for a more nuanced capture of the underlying dynamics in volatility surfaces. Recent studies have highlighted the promise of GPs in volatility modeling and option pricing (Qin and Almeida 2020; Tegner and Roberts 2021). Nonetheless, significant questions still exist regarding the refinement of GP implementations in this domain.

Building upon these promising developments, our research looks into enhancing Gaussian process models specifically tailored for options. We analyze the effectiveness of various kernel functions and model architectures in capturing hidden relations within implied volatility. Furthermore, we assess the performance of our models in data-sparse environments. In all cases, we utilize our methodology to generate price confidence intervals that appropriately quantify the inherent uncertainties in the modeling process, a critical feature for risk management. In aggregate, our work thoroughly examines the viability of Gaussian processes as a robust and reliable approach to index option pricing.

3 Literature Review

Options are contracts between the buyer and seller that give the purchaser the right to buy (call options) or sell (put options) an underlying asset at a predetermined time and price. To model the stock price S, we utilize Geometric Brownian Motion, defined by the stochastic differential equation:

$$dS_t = rS_t dt + \sigma(t, S_t) S_t dW_t \tag{1}$$

where r represents the risk-free rate and $\sigma(t, S_t)$ is the volatility, a deterministic function of the time t as well as the stock price S_t.

For a European option with strike Price K and time to expiry of T, The Black-Scholes model (Black and Scholes 1973) provides a pricing formula for call and put options:

$$C(t, S_t, T, K) = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{2}$$

$$P(t, S_t, T, K) = Ke^{-r(T-t)} - \Phi(-d_2) - S_t\Phi(-d_1) \tag{3}$$

Where:

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left[\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right]$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

Since the introduction of the Black-Scholes formulas, there have been key advances in volatility models, namely local and stochastic volatility. Heston first introduced the stochastic volatility model (Heston 1993), where volatility itself follows a stochastic process:

$$d\sigma_t = a(b - \sigma_t)dt + \xi\sqrt{\sigma_t}dW_t \tag{4}$$

Dupire then introduced the local volatility model (Dupire 1994), where volatility is treated as a deterministic function of $\sigma(T, K)$, governed by Dupire's equation:

$$\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K} - \frac{K^2\sigma^2(T,K)}{2}\frac{\partial^2 C}{\partial K^2} = 0 \tag{5}$$

These two models can provide well-performing solutions for the implied volatility problem. However, they may rely on unstable or complex calibration methods to get to that solution, requiring numerical integration or other computations.

Dumas et al. proposed a simpler polynomial ad hoc approximation for estimating implied volatility (Dumas, Fleming, and Whaley 1998). This approach uses a polynomial function to approximate the surface:

$$\sigma(T, K) = \beta_0 + \beta_1 T + \beta_2 T^2 + \beta_3 K + \beta_4 K^2 + \beta_5 TK \tag{6}$$

Furthermore, there are machine-learning approaches to option pricing. Hutchinson et al. proposed a neural-network-based pricing and hedging method (Hutchinson, Lo, and Paggio 1994), and more recently, Buehler et al. suggested a reinforcement-learning approach to option pricing (Buehler et al. 2019). These methods, however, give point price estimates and are extremely data-intensive.

Finally, the two papers introduce GP implementations for solving the option pricing problem. First, Tegnér and Roberts used a Squared Exponential Kernel on the implied volatility and then a Monte Carlo method to fit market prices (Tegner and Roberts 2019). Then, Qin and Almeida used a variety of different Kernels to determine the best-performing method (Qin and Almeida 2020). Our work aims to build on these two papers to further demonstrate the applicability of GPs in this domain.

4 Data & Preprocessing

Our database is split into 2 sections; first, we worked on statistical analysis and model development using a small testing set, of 36,604 observations (only a month's worth of option chain). Due to the complexity, train time, and size of our models, we found it fit to implement our initial research on a robust and easily accessible dataset, through which we would confirm any preliminary relationships and phenomena between aspects of the data and our model outputs. After thorough testing and validation, during which we narrowed down our model's parameters, we moved forward with our full dataset to hyper-tune our model of choice and improve the accuracy and smoothness of our implied volatility surface. While these 2 datasets are mutually exclusive, the necessary relationships and factors are common between the 2 datasets to retain consistency in our models.

4.1 Testing Dataset

Given that the S&P 500 options are among the most liquid and commonly studied, we aimed to apply our Bayesian models on the S&P 500 Exchange-Traded Fund (SPY) American options option chain for our estimation.

We utilized the Option EOD Summary product from CBOE DataShop [1] to access SPY option data for the full month of October 2023 (10/2/23 - 10/31/23) as our initial dataset. The data for each option includes its quote date, expiration date, strike price, option type (call vs put), close price, active underlying price, implied volatility, end-of-day (EOD) bid/ask values, and sizes. CBOE calculated the implied volatility for the options using a binomial asset pricing model with discretized dividends. Data from Bloomberg, Polygon, and other sources are similarly sufficient as well, which we are continuing to experiment with to enhance the accuracy and precision of our models under larger datasets.

Table 1 contains a summary of the preliminary dataset we are using, spanning across the 22 trading days in October 2023. There are some trends apparent in the data that are important to conclude from, formed by the deconstruction of our data into 9 distinct categories. Our options have been classified into 3 distinct time-to-maturities, with anything less than 20 days being short-term and anything longer than 365 days being long-term. Our at-the-money (ATM) options are of moneyness between 0.99 and 1.01. In-the-money (ITM) options and out-the-money (OTM) options for calls are defined as moneyness greater than 1.01 and moneyness less than 0.99 respectively, and the opposite is true for put options.

Table 1: **Testing Data Summary**

			Total	**ATM**	**ITM**	**OTM**
call options	data count	**Total**	36604	4341	10055	22208
		1-20 days	11601	1822	3280	6499
		20-365 days	19156	2034	4859	12263
		365-1095 days	3475	151	961	2363
	average implied volatility	**Total**	26.13%	18.13%	40.97%	20.97%
		1-20 days	19.52%	16.64%	25.89%	17.11%
		20-365 days	16.86%	16.96%	22.91%	14.45%
		365-1095 days	17.95%	19.43%	23.62%	15.55%
	average option price	**Total**	$15.96	$11.81	$42.44	$4.77
		1-20 days	$7.29	$4.71	$22.25	$0.46
		20-365 days	$16.83	$16.74	$47.24	$4.80
		365-1095 days	$43.48	$53.98	$102.85	$18.67
put options	data count	**Total**	44309	4427	11084	28798
		1-20 days	15115	1839	3527	9749
		20-365 days	22985	2093	5943	14949
		365-1095 days	3467	160	857	2450
	average implied volatility	**Total**	33.25%	18.19%	30.84%	36.50%
		1-20 days	27.36%	16.65%	21.00%	31.68%
		20-365 days	24.67%	17.01%	17.57%	28.57%
		365-1095 days	24.89%	19.47%	19.91%	26.99%
	average option price	**Total**	$10.26	$8.69	$29.24	$3.20
		1-20 days	$5.10	$4.13	$18.29	$0.51
		20-365 days	$11.67	$12.16	$30.81	$3.99
		365-1095 days	$27.22	$30.78	$72.18	$11.26

First, we can establish the presence of a robust volatility smile since, on average, the implied volatilities

1. https://datashop.cboe.com

of ATM options are less than the ITM and out-the-money OTM options. We see OTM options are far cheaper, while ITM options are far more expensive than ATM options. Additionally, options with a longer time-to-maturity are priced higher than those expiring soon due to the time value of money. Something to consider is that there are more total puts than calls in the dataset, which may influence model results.

We have also imposed several filters and conditions upon our dataset to motivate precision in our models. Options with a trading volume of 0 have been dropped from our dataset to remove options that have not been traded, likely due to mispricing or lack of liquidity. Options with an implied volatility within 0 (non-inclusive) and 1 (inclusive) were maintained in our dataset, while we considered any others as outliers. Further preprocessing included constraining the moneyness of our options to within 0.7 and 1.3, as well as our time-to-maturity to within 20-365 days. This allowed us to further filter out the possibility of inconsistent results due to outliers by excluding deep ITM and OTM options, as well as options that expire too soon and too far in the future. After all preprocessing and cleaning, our dataset was cut in half from $\approx$ 81,000 initial observations to $\approx$ 40,000 unique data points.

4.2 Final Dataset

Several changes were made from our first dataset to our final dataset based on numerous obstacles and complications met during our testing period. First, we altered our underlying asset from the SPY to the S&P 500 Index (SPX), meaning we could exclude factors such as expense ratio and tracking error between SPY and SPX, and more importantly, use less complex European option pricing models with our results. The downside of using SPX over SPY is that SPY is a more convenient investment opportunity to retail investors than SPY is, but with both assets being largely traded, the difference in liquidity is negligible. Rather than 22 trading days, we opted to query 3 years of post-COVID option chain between 2021-01-01 to 2024-01-01. Along with the option chain data, we also obtained E-mini S&P 500 Index Futures over the same time period. This data allowed us to refine our moneyness calculations and interpolate the implied yield curve for our option prices.

While there were several obstacles in obtaining the data, with attempted queries from several sources including Bloomberg, our final dataset was pulled from Refinitiv's Elektron time series set, provided by the Hanlon Financial Systems Laboratory at Stevens Institute of Technology. The data consisted of just under 5 million observations and many useful factors. Some of the significant inputs include option type, bid-ask, expiration date, underlying price, and last trade price. Refinitiv calculated the implied volatility for the European options through an inversion of the Black-Scholes option pricing model, which proved useful as a comparison to our point estimates. Our moneyness has been defined under the same boundaries as in our testing dataset, only now using a forward-looking price for the underlying rather than the immediate spot price.

Table 2 displays similar summary statistics for our final dataset, this time consisting of the historical SPX option chain from January 2021 to December 2023. It is important to note that the option prices are much larger in magnitude compared to the SPY options largely due to the underlying assets' difference in price; specifically, SPX trades $\approx$ 10 times higher than SPY. Through these summary statistics, we can confidently establish the robust trends we saw in our testing dataset, including the existence of a volatility smile and the time value of money. These relationships are important as they influence the structure of implied volatility surfaces, of which the axes are defined by moneyness and maturity.

During our thorough data cleaning, we were able to filter our selection of contracts under similar conditions as in our testing dataset. By including only observations that range between 0.7 to 1.3 moneyness and 20 to 365 days to maturity, we aim to remove potential bias and variability from abnormal and illiquid contracts. Observations with missing values and 0 trading volume were also removed for the same reasons. Post-processing, our inputs consisted of $\approx$ 3.43 million observations as calculated by the sum of total calls and puts, which is almost 85 times as much input data as in our preliminary models.

Table 2: **Final Data Summary**

			Total	**ATM**	**ITM**	**OTM**
call options	data count	**Total**	2254487	85469	855467	686136
		1-20 days	161020	7570	72667	53508
		20-365 days	1673637	67657	670161	522670
		365-1095 days	344229	9049	101299	97557
	average implied volatility	**Total**	27.98%	17.97%	26.03%	17.41%
		1-20 days	53.95%	17.1%	41.58%	25.61%
		20-365 days	26.16%	17.94%	24.23%	16.39%
		365-1095 days	21.87%	19.3%	22.18%	16.71%
	average option price	**Total**	$628.16	$188.07	$568.14	$73.94
		1-20 days	$535.78	$46.20	$473.01	$4.52
		20-365 days	$606.47	$174.57	$552.86	$52.66
		365-1095 days	$770.35	$402.34	$738.82	$183.05
put options	data count	**Total**	2395408	85467	654872	861465
		1-20 days	169014	7558	47353	75347
		20-365 days	1754289	67645	499880	671238
		365-1095 days	379176	9046	96220	101401
	average implied volatility	**Total**	32.32%	18.72%	20.79%	26.79%
		1-20 days	52.1%	16.59%	35.08%	36.9%
		20-365 days	30.82%	18.69%	19.06%	25.63%
		365-1095 days	28.51%	20.97%	18.84%	24.39%
	average option price	**Total**	$321.46	$192.67	$594.01	$99.22
		1-20 days	$200.22	$47.87	$465.89	$5.51
		20-365 days	$268.53	$176.44	$559.99	$79.70
		365-1095 days	$473.81	$404.72	$785.09	$255.67

5 Methodology

5.1 Gaussian Process Formulation

Let us denote the implied volatility of a given option as a process $\sigma(x_i)$, where $x_i = (M_i, T_i)$. M_i denotes the moneyness of the option, which we calculate as its future value divided by its strike price. T_i is the time to maturity of the option, measured in days. In our temporal models, we include days since the quote occurred, denoted as D_i, making $x_i = (M_i, T_i, D_i)$. Furthermore, we assume that implied volatility, σ, is a continuous positive value function. Additionally, we log transform σ as $f(x) = \ln(\sigma(x))$ to improve the predictive performance of our model.

The prior distribution of processes is entirely described by a mean function, $m(x)$, and a covariance (kernel) function, $k(x, x')$, defined as:

$$\begin{aligned} m(x) &= \mathbb{E}[f(x)] \\ k(x, x') &= \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))] \end{aligned}$$

The Gaussian process then generates the posterior distribution of functions as follows:

$$f(x) \sim \mathbb{GP}(m(x), k(x, x'))$$

It is important to note that this process is categorized as non-parametric because it will define the function space with infinite parameters. For each $x \in \mathbb{R}$, there will be a parameterized posterior distribution.

Here, the kernel function measures the similarity between the data points across a given process and is typically a symmetric function based on a distance metric. The kernel can be formulated in different ways, which would result in differing prior distributions of processes on which to perform Bayesian inference. The kernel we employ is the squared exponential kernel, formulated as:

$$k_{SE} = \exp(\frac{r^2}{2})$$

where r is parameterized as

$$r = \sqrt{\frac{(M - M')^2}{2l_M^2} + \frac{(T - T')^2}{2l_T^2}}$$

and l_M and l_T are the kernel hyper-parameters to be optimized on.

5.2 Model Training Specification

The steps in training our Exact Gaussian Process are explicitly described in Gaussian Processes for Machine Learning (Rasmussen and I. 2008) by Rasmussen and Williams. We first start by randomly initializing our kernel hyperparameters. Then, in each subsequent iteration, we compute the mean and covariance matrix, as well as its inverse and determinant. This is the root of the computational complexity involved with this method. With these calculations, we can derive the log marginal likelihood, which is the loss function used for the Exact Gaussian Process. The likelihood function measures how well our model explains the observed data under its current state. Finally, to optimize the hyperparameters of the kernel function, we calculate the gradient of the log marginal likelihood with respect to these parameters. Using the Adaptive Moment Estimation (Adam) optimization algorithm, we minimize the log marginal likelihood to find the best-fitting model. These steps are repeated until the optimization converges.

6 Model Setup

Our first model imitates the model found in Qin and Almeida's research. This consists of two exact inference models trained on a single day of option quotes - one for calls and one for puts. Since the number of observations is relatively small, the exact inference method can be used, wherein we reserve 20% of data for testing. We stratify the testing space over moneyness into three strata per time to maturity, in order to maintain an equal distribution of in-the-money to out-the-money option quotes in the training process. This is because these quotes are relatively sparse and the majority of the data density is closer to the at-the-money quotes. This improves our out-of-sample performance, which we evaluate with root-mean-squared-error (RMSE). As a baseline, we also estimate the implied volatility surface each day with the ad-hoc solution provided by Dumas (Dumas, Fleming, and Whaley 1998).

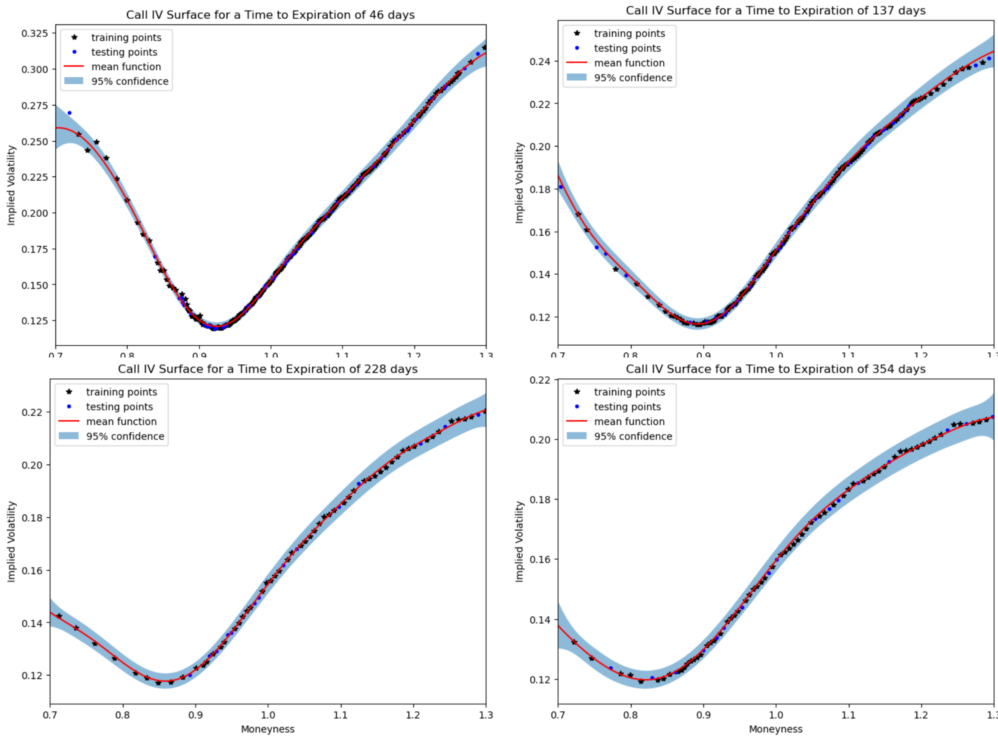
Furthermore, we use the generated implied volatility confidence intervals to price options in order to compare them to the market prices. We use a simple Black-Scholes model approach, bootstrapping the risk-free-rate from daily treasury par yield curve rates. Using these estimates, we calculate the mean valuation error (MVE) and the mean percent error (MPE) as our measures of pricing efficiency. We can also find the percentage of quotes that land in the bid-ask spread. Using these measures, we can model arbitrage opportunities.

7 Results

7.1 Implied Volatility Confidence Intervals

Using October 2nd, 2023 as an arbitrarily chosen date, we display cross-sections of the interpolated implied volatility surface from our model. In these visualizations, we show the confidence intervals for each unique moneyness, time to expiration pair, as well as the training and testing points for each.

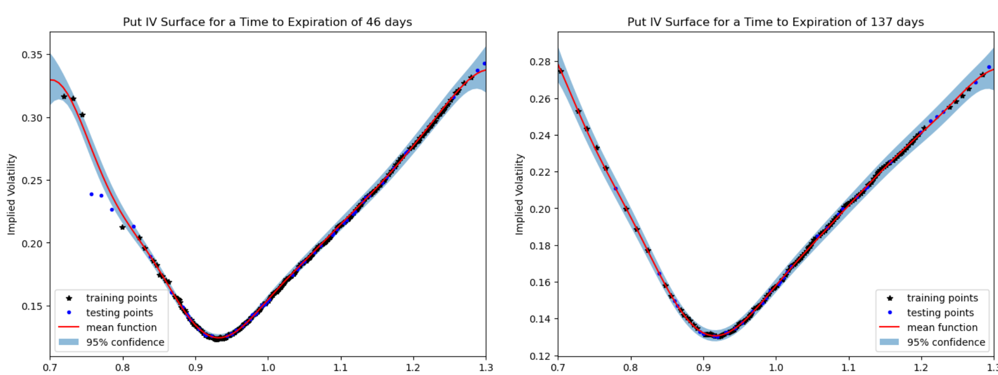
Calls

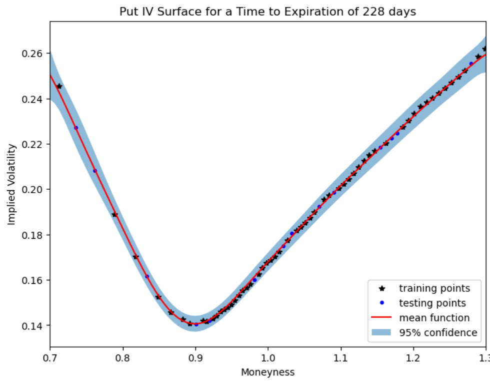


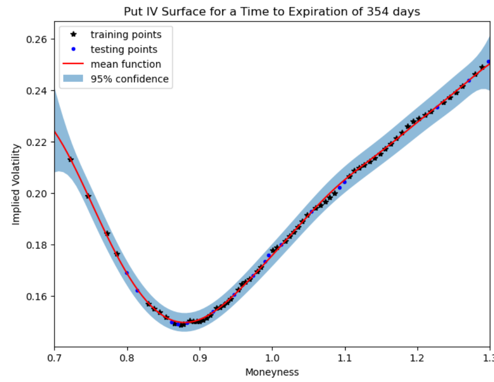
Our call model displays the characteristic volatility smile through the mean function, as well as flattening across further maturities. As for the confidence interval, we see a widening across further maturities as well, as well as deep ITM and OTM options. This can be attributed to data sparsity. The closer maturities have more actively traded strikes, and therefore the confidence intervals are much tighter. The same effect was observed for deep ITM and OTM options. These options are less frequently traded, and therefore the confidence interval on implied volatility widens in these areas.

We can also see an artifact of noise within our in-sample implied volatilities. Since these are based on market quotes, the variation in these quotes results in noisy implied volatilities that do not conform to our implied volatility surface. This can be especially seen in the calls that have moneyness less than 1. Theoretically, these implied volatilities would imply that these options are mispriced. However, due to their calculation from the midprice, these trades would not be able to be executed at a beneficial price.

Puts



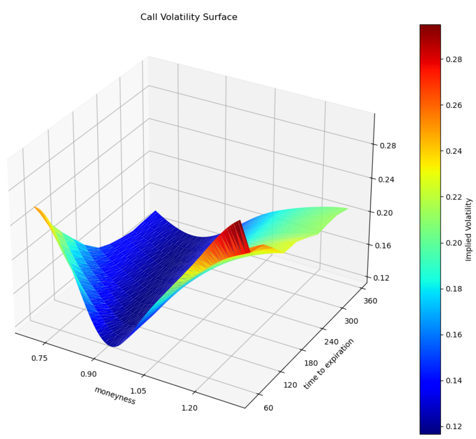


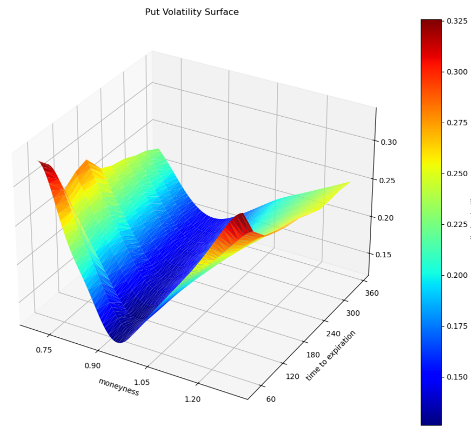


Our put models demonstrate similar aspects to the call models, namely the widening of confidence intervals in characteristic locations and flattening across further maturities. Another trait of the model setup can be seen at the edges of the moneyness aspect. Since we constrain our option quotes from 0.7 and 1.3 moneyness, this results in limited training data around the edges, resulting in a widening of the curve. If we were to extend the training range across moneyness, this effect might be lessened.

7.2 Volatility Surfaces

The entire volatility surface for the call and put model is generated as well, using the mean function of the Gaussian Process.





We find a smooth implied volatility surface generated from our models, showing the characteristic flattening of implied volatilites with further maturities. From this, we conclude our model achieves a good fit that captures the underlying patterns well.

7.3 Evaluation

7.3.1 Implied Volatility Results

We first evaluate the models on their fit to implied volatility and a comparison to the ad-hoc Dumas approximation. We see a distinct out-performance of the ad-hoc approximation overall, and overall moneyness for both calls and puts.

Specifically, we see a decrease in performance for the in-the-money portions of both the call and put models. This can be attributed to the noise seen in the previous cross-sections of the implied volatility surfaces, in which noisy quotes provide noisy estimates of implied volatility.

RMSE	Model	Dumas
Overall	0.004059	0.022226
OTM	0.002748	0.025278
ATM	0.000834	0.014365
ITM	0.004462	0.019854

Table 3: Calls RMSE Evaluation

RMSE	Model	Dumas
Overall	0.003799	0.023898
OTM	0.001566	0.018374
ATM	0.000898	0.016576
ITM	0.005541	0.030018

Table 4: Puts RMSE Evaluation

7.3.2 Pricing Results

We also compare the market mid-prices with the prices generated from our models using a Black-Scholes model. From this, we find the overall mean valuation error to be around eight dollars. Put into more perspective, we find the mean percent error to be around 5% for calls and 4% for puts.

	Calls	**Puts**
Mean Value Error (MVE)		
Overall	8.078337	7.876176
Out of The Money (OTM)	3.464030	3.700192
At The Money (ATM)	7.809565	7.068187
In The Money (ITM)	11.643077	13.594968
Percent Error		
Overall	0.051903	0.038585
Out of The Money (OTM)	0.095105	0.046215
At The Money (ATM)	0.040569	0.036521
In The Money (ITM)	0.022507	0.027937

Table 5: Detailed Statistical Measures for Calls and Puts

The highest mean valuation error is for in-the-money options. However this is due to the dynamics of option prices themselves, as we can see that percentage-wise, these options are priced closest to their mid-price.

8 Conclusion

With our results, we observe certain behaviors of our model that we sought after. First and most importantly, we find posterior distributions for the implied volatility function that provide much more information than a point-wise estimate. With these posterior distributions, we can calculate confidence intervals as seen before and probabilistic ally quantify market mispricings according to our model. Outperforming traditional parametric benchmarks such as the Dumas polynomial approximation, our research substantiates that GPs can provide a more dynamic and informative framework, particularly beneficial for the nuanced domain of options pricing.

References

Black, Fischer, and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (3): 637–654. ISSN: 00223808, 1537534X, accessed November 27, 2023. http://www.jstor.org/stable/1831029.

Buehler, Hans, Lukas Gonon, Josef Teichmann, Ben Wood, Baranidharan Mohan, and Jonathan Kochems. 2019. Deep hedging: hedging derivatives under generic market frictions using reinforcement learning. *SSRN Electronic Journal* (May). https://doi.org/10.2139/ssrn.3355706.

Dumas, Bernard, Jeff Fleming, and Robert E. Whaley. 1998. Implied volatility functions: empirical tests. *The Journal of Finance* 53 (6): 2059–2106. ISSN: 00221082, 15406261, accessed April 24, 2024. http://www.jstor.org/stable/117461.

Dupire, Bruno. 1994. Pricing with a smile (January). https://www.risk.net/derivatives/equity-derivatives/1500211/pricing-with-a-smile.

Heston, Steven L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6 (2): 327–343. ISSN: 08939454, 14657368, accessed November 27, 2023. http://www.jstor.org/stable/2962057.

Hutchinson, James, Andrew Lo, and Tomaso Paggio. 1994. A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance* 49 (3): 851–889. https://doi.org/https://doi.org/10.1111/j.1540-6261.1994.tb00081.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1994.tb00081.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1994.tb00081.x.

Qin, Zhang, and Caio Almeida. 2020. A bayesian nonparametric approach to option pricing. *Brazilian Review of Finance* 18, no. 4 (November): 115–137. https://doi.org/10.12660/rbfin.v18n4.2020.81913.

Rasmussen, Carl Edward, and Williams Christopher K I. 2008. *Gaussian processes for machine learning.* MIT Press.

Tegner, Martin, and Stephen Roberts. 2019. A Probabilistic Approach to Nonparametric Local Volatility, no. 1901.06021 (January). https://ideas.repec.org/p/arx/papers/1901.06021.html.

———. 2021. A Bayesian take on option pricing with Gaussian processes, no. 2112.03718 (December). https://ideas.repec.org/p/arx/papers/2112.03718.html.