

Fitting Volatility Smiles with Structural Models and Machine Learning

Edward Parcell

March 16, 2026

Abstract

We propose a hybrid method for interpolating and extrapolating the implied volatility smile that combines the SABR stochastic volatility model with Gaussian Process regression. Rather than using a Gaussian Process to model Black-Scholes implied volatility directly as a function of strike, we first calibrate SABR globally to the observed smile, then invert the SABR formula at each market strike to find the value of the SABR initial volatility parameter σ_0 that reproduces the market quote. The resulting function $\sigma_0(K)$ is nearly flat when SABR provides a good structural fit, making it well-suited to Gaussian Process interpolation. The method inherits SABR's economically motivated wing behavior for extrapolation while achieving exact calibration at observed strikes, and provides natural uncertainty quantification through the Gaussian Process posterior.

1 Introduction

The implied volatility smile is a fundamental object in options markets. Since its emergence after Black Monday in 1987, practitioners have needed methods to interpolate the smile between observed market strikes and, critically, to extrapolate it into the wings where liquid quotes may be sparse or absent.

The tension in smile fitting is between flexibility and structure. Purely parametric models such as SABR (Hagan et al., 2002) or SVI (Gatheral, 2004) impose a functional form that provides smooth interpolation and controlled extrapolation, but may not fit all market quotes exactly. Purely nonparametric methods such as splines or Gaussian Processes can fit observed data precisely, but may produce implausible behavior between and beyond observed strikes, and particularly in the wings, where data is sparse and extrapolation is most consequential.

In this paper we propose a hybrid approach that combines the strengths of both. We use SABR as a structural backbone to capture the overall shape of the smile, then apply Gaussian Process regression to model the residual discrepancy between SABR and the market. The key insight is in the choice of what to model with the Gaussian Process: rather than modeling Black-Scholes implied volatility $\sigma_{BS}(K)$ directly, we let the SABR initial volatility parameter σ_0 vary as a function of strike, and model this function $\sigma_0(K)$ with the GP. When SABR

provides a reasonable global fit, $\sigma_0(K)$ is nearly flat, making it an ideal target for GP regression.

The advantages of this decomposition are threefold. First, interpolation between market strikes is handled by a GP operating on a smooth, slowly varying function, which is precisely the regime in which GP regression excels. Second, extrapolation beyond observed strikes is governed by SABR: as the GP posterior mean reverts to the global best-fit σ_0 in the absence of nearby data, the resulting smile inherits SABR’s dynamically motivated wing behavior. Third, the GP posterior variance provides a natural measure of model uncertainty that increases with distance from observed data.

The remainder of this article is organized as follows. Section 2 recaps the volatility models used in this work. Section 3 provides a brief introduction to Gaussian Process regression. Section 4 describes our fitting method. Section 5 presents results. Finally, Section 6 gives conclusions.

2 Recap of Volatility Models

2.1 Black-Scholes Model

In the classic Black-Scholes model for valuing options, the price of an underlying is taken to evolve according to a geometric Brownian motion process:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{1}$$

By showing that an instantaneously hedged portfolio can be constructed using the option and the underlying, and that this portfolio must return the risk-free rate, Black, Scholes and Merton were able to construct an explicit formula for the price of European call and put options:

$$V^C = \Phi(d_1)S - \Phi(d_2)Ke^{-rt} \tag{2}$$

$$V^P = \Phi(-d_2)Ke^{-rt} - \Phi(-d_1)S \tag{3}$$

$$d_{1,2} = \frac{\ln(S/K) + (r \pm \frac{1}{2}\sigma^2)t}{\sigma\sqrt{t}} \tag{4}$$

The drift of the stock price does not appear in this solution, and so Black-Scholes establishes a correspondence between option price and volatility. Because of this, volatility is used as a standard communication tool in equity, futures, interest rate, and FX option markets.

Prior to 1987, option prices with different strikes tended to have prices consistent with the same volatility. After Black Monday showed the risk associated with deep out-of-the-money puts, the volatility smile emerged, where options away from the at-the-money strike tend to price with higher volatilities.

2.2 SABR Model

In the SABR model (Hagan et al., 2002), instantaneous volatility is taken to be a process that itself evolves according to a second, correlated Brownian motion process:

$$dF_t = \sigma_t(F_t)^\beta dW_t \quad (5)$$

$$d\sigma_t = \alpha\sigma_t dZ_t \quad (6)$$

$$dW_t dZ_t = \rho dt \quad (7)$$

where F_t is the forward price, σ_t is the instantaneous volatility with initial value σ_0 , α is the volatility of volatility, $\beta \in [0, 1]$ is the CEV exponent, and $\rho \in [-1, 1]$ is the correlation between the forward and its volatility.

Hagan et al. computed an asymptotic expansion for Black-Scholes implied volatility σ_{impl} as a function of $\varepsilon = T\alpha^2$.

$$\sigma_{\text{impl}} = \alpha \frac{\log(F_0/K)}{D(\zeta)} \left\{ 1 + \left[\frac{2\gamma_2 - \gamma_1^2 + 1/(F_{\text{mid}})^2}{24} \left(\frac{\sigma_0 C(F_{\text{mid}})}{\alpha} \right)^2 + \frac{\rho\gamma_1}{4} \frac{\sigma_0 C(F_{\text{mid}})}{\alpha} + \frac{2 - 3\rho^2}{24} \right] \varepsilon \right\}, \quad (8)$$

where, $C(F) = F^\beta$ and $F_{\text{mid}} = \frac{F_0+K}{2}$ and

$$\zeta = \frac{\alpha}{\sigma_0} \int_K^{F_0} \frac{dx}{C(x)} = \frac{\alpha}{\sigma_0(1-\beta)} (F_0^{1-\beta} - K^{1-\beta}), \quad (9)$$

$$\gamma_1 = \frac{C'(F_{\text{mid}})}{C(F_{\text{mid}})} = \frac{\beta}{F_{\text{mid}}} \quad (10)$$

$$\gamma_2 = \frac{C''(F_{\text{mid}})}{C(F_{\text{mid}})} = -\frac{\beta(1-\beta)}{(F_{\text{mid}})^2}, \quad (11)$$

$$D(\zeta) = \log \left(\frac{\sqrt{1 - 2\rho\zeta + \zeta^2} + \zeta - \rho}{1 - \rho} \right). \quad (12)$$

In practice, one of the major advantages of the SABR model is that its closed-form approximation can quickly be calibrated to a market smile using a multivariate optimizer. This is done by attempting to minimize some measure of error between Black-Scholes implied volatilities calculated using the SABR model, $\sigma_{\text{impl}}(K)$, and Black-Scholes implied volatilities calculated directly from market prices, $\sigma_{\text{mkt}}(K)$. Using the ℓ_2 norm for example, we choose σ_0 , α , β and ρ to minimize

$$\sum_i (\sigma_{\text{impl}}(K_i) - \sigma_{\text{mkt}}(K_i))^2 \quad (13)$$

Typically the problem is close to degenerate as changes to β and ρ have a similar effect on the smile, so β is usually held constant. Additionally, the asymptotic expansion is not scale-free, so in some contexts it is usual to rescale strikes so that the ATM strike is 1.

The crucial feature of SABR for our purposes is the smile shape it produces: the skew and curvature arise from the stochastic evolution of volatility and its correlation with the underlying, not from an arbitrary functional form. This gives economic meaning to the extrapolation behavior in the wings.

3 Gaussian Processes

A Gaussian Process (GP) is a distribution over functions such that any finite collection of function values is jointly Gaussian (Rasmussen and Williams, 2006). A GP is fully specified by a mean function $m(x)$ and a covariance (kernel) function $k(x, x')$:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \quad (14)$$

Given n observations $\mathbf{y} = [y_1, \dots, y_n]^\top$ at inputs $\mathbf{X} = [x_1, \dots, x_n]^\top$ with i.i.d. Gaussian noise of variance σ_n^2 , the posterior predictive distribution at a test point x_* is Gaussian with mean and variance:

$$\mu(x_*) = m(x_*) + \mathbf{k}_*^\top (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} (\mathbf{y} - \mathbf{m}) \quad (15)$$

$$\sigma^2(x_*) = k(x_*, x_*) - \mathbf{k}_*^\top (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}_* \quad (16)$$

where \mathbf{K} is the $n \times n$ matrix with entries $K_{ij} = k(x_i, x_j)$, \mathbf{k}_* is the n -vector with entries $k(x_i, x_*)$, and \mathbf{m} is the n -vector of prior mean evaluations.

What this means in practice is that Gaussian Processes provide a naturally smooth interpolator given a set of control points that map from \mathbb{R}^n to \mathbb{R} .

Several properties make GPs well-suited to our application:

- **Mean reversion.** Far from observed data, the posterior mean reverts to the prior mean function $m(x)$. In our setting, the mean function is $\hat{\sigma}_0$ everywhere, and this means the model reverts to the globally calibrated SABR parameters in the wings.
- **Uncertainty quantification.** The variance estimate in (16) increases with distance from observations, providing a natural measure of interpolation and extrapolation uncertainty.
- **Smoothness control.** The choice of kernel controls the smoothness of the interpolated function. We use the squared exponential (RBF) kernel

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{(x - x')^2}{2\ell^2}\right) \quad (17)$$

where σ_f^2 is the signal variance and ℓ is the length scale.

- **Interpolation can be exact, or incorporate uncertainty.** If we wish our interpolation to precisely hit its control points, we can let $\sigma_n^2 \rightarrow 0$ and the GP posterior passes exactly through the observed data points. Alternatively, we can use the bid-ask spread, for example, to allow for uncertainty, even at control points.

When used generally, the kernel hyperparameters σ_f , ℓ , and σ_n are typically estimated by maximizing the log marginal likelihood:

$$\log p(\mathbf{y} \mid \mathbf{X}, \theta) = -\frac{1}{2} \mathbf{y}^\top \mathbf{K}_y^{-1} \mathbf{y} - \frac{1}{2} \log |\mathbf{K}_y| - \frac{n}{2} \log 2\pi \quad (18)$$

where $\mathbf{K}_y = \mathbf{K} + \sigma_n^2 \mathbf{I}$ and θ collects the hyperparameters.

However, for dealing with volatility smiles, it is helpful to either fix or constrain hyperparameters, to ensure consistency in smiles fitted at different maturities and on different days.

4 Fitting Method: SABR + Gaussian Processes

We now describe the proposed hybrid fitting method. Suppose we observe market implied volatilities $\sigma_{\text{mkt}}(K_i)$ at a discrete set of strikes K_1, \dots, K_n for a given expiry T and forward price F_0 .

4.1 Step 1: Global SABR Calibration

We first calibrate SABR globally to the observed smile. Let us write $\mathcal{S}(K_i; \sigma_0, \alpha, \beta, \rho, F_0, T)$ for σ_{impl} in (8). Fixing β (typically to a market-conventional value such as $\beta = 0.5$ or $\beta = 1$), we find $\hat{\sigma}_0$, $\hat{\alpha}$, and $\hat{\rho}$ by minimizing

$$\sum_{i=1}^n (\mathcal{S}(K_i; \sigma_0, \alpha, \beta, \rho, F_0, T) - \sigma_{\text{mkt}}(K_i))^2 \quad (19)$$

This yields a best-fit SABR smile $\hat{\sigma}_{\text{SABR}}(K) = \mathcal{S}(K; \hat{\sigma}_0, \hat{\alpha}, \beta, \hat{\rho}, F_0, T)$ that captures the broad shape of the market smile but will generally not pass through every market quote exactly.

4.2 Step 2: Strike-Dependent Inversion

At each observed strike K_i , we invert the SABR formula to find the value $\sigma_0^{(i)}$ such that SABR *exactly* reproduces the market implied volatility:

$$\mathcal{S}(K_i; \sigma_0^{(i)}, \hat{\alpha}, \beta, \hat{\rho}, F_0, T) = \sigma_{\text{mkt}}(K_i) \quad (20)$$

This is a one-dimensional root-finding problem for each strike, which can efficiently and robustly be solved using standard numerical solvers. The other SABR parameters ($\hat{\alpha}, \beta, \hat{\rho}$) are held at their previously calibrated values.

If SABR provides a good global fit, the values $\sigma_0^{(i)}$ will be close to the global best-fit $\hat{\sigma}_0$, and the function $\sigma_0(K)$ will be nearly flat. This is the key property that makes the GP interpolation well-behaved.

4.3 Step 3: Gaussian Process Interpolation

We model $\sigma_0(K)$ as a Gaussian Process with prior mean equal to the global SABR calibration:

$$\sigma_0(K) \sim \mathcal{GP}(\hat{\sigma}_0, k(K, K')) \quad (21)$$

where the constant prior mean $m(K) = \hat{\sigma}_0$ ensures that far from observed data, the GP posterior reverts to the globally calibrated value. We use the squared exponential kernel (17), with hyperparameters estimated by maximizing the marginal likelihood on the observations $\{(K_i, \sigma_0^{(i)})\}_{i=1}^n$.

We can set the noise variance σ_n^2 to a very small value (or zero) since the inversion in Step 2 is exact: the $\sigma_0^{(i)}$ are not noisy observations of an underlying function, but computed values. Alternatively, the noise variance can be set to reflect uncertainty using the difference between volatilities implied from bid and ask prices.

4.4 Step 4: Evaluation

To compute the model-implied volatility at an arbitrary strike K_* , we:

1. Evaluate the GP posterior mean: $\tilde{\sigma}_0(K_*) = \mu_{\text{GP}}(K_*)$ using (15).
2. Evaluate the SABR formula with this strike-dependent initial volatility:

$$\sigma_{\text{model}}(K_*) = \mathcal{S}(K_*; \tilde{\sigma}_0(K_*), \hat{\alpha}, \beta, \hat{\rho}, F_0, T) \quad (22)$$

The GP posterior variance $\sigma_{\text{GP}}^2(K_*)$ provides an estimate of model uncertainty, which can be propagated through the SABR formula to obtain confidence bands on the implied volatility smile.

4.5 Properties of the Method

Exact calibration is possible. In the noise-free case, at each observed strike K_i the GP posterior mean returns $\sigma_0^{(i)}$, so $\sigma_{\text{model}}(K_i) = \sigma_{\text{mkt}}(K_i)$ by construction.

Structural extrapolation. For strikes far from any observation, $\tilde{\sigma}_0(K_*) \rightarrow \hat{\sigma}_0$, and the model reduces to the globally calibrated SABR smile. The wing behavior is therefore governed by SABR’s dynamically motivated smile shape rather than an arbitrary extrapolation rule.

Smooth interpolation. Between observed strikes, $\sigma_0(K)$ varies smoothly (governed by the GP kernel), producing a smooth implied volatility smile. Since the GP is interpolating a nearly flat function, the risk of overfitting or oscillation is minimal.

Comparison with direct GP on $\sigma_{\text{BS}}(K)$. One could instead apply a GP directly to Black-Scholes implied volatilities. This is simpler but suffers from two drawbacks. First, $\sigma_{\text{BS}}(K)$ has significant curvature, making the GP’s job harder and increasing the risk of implausible wiggles between observations. Second, extrapolation in the wings reverts to a constant (the GP prior mean), which produces implausible flat wings. By contrast, our method’s wing extrapolation inherits the skew and curvature of SABR.

4.6 Practical Considerations

Choice of kernel. The squared exponential kernel is a natural default, as it produces infinitely differentiable functions and has only two free hyperparameters. The Matérn- $\frac{5}{2}$ kernel is a reasonable alternative that produces smoother-than-spline but not infinitely smooth interpolants. In our experiments, both produce similar results.

Breakdown of the SABR expansion. The Hagan asymptotic expansion is known to break down for long maturities or strikes far from at-the-money, where the expansion parameter $\varepsilon = T\alpha^2$ is not small. However, the important factor remains the economically plausible shape provided by the SABR model.

Arbitrage freedom. The method does not explicitly enforce no-arbitrage constraints (non-negative butterfly spreads, monotone call prices). In practice, when SABR provides a reasonable global fit so that $\sigma_0(K)$ is nearly flat, arbitrage violations are unlikely. However, for production use, a post-hoc check and, if necessary, a constrained GP formulation (?) may be advisable.

Extension to the surface. For a single expiry, the method operates as described. To construct a full implied volatility surface, one can apply the method independently at each expiry (since SABR is calibrated per-expiry) and then interpolate the GP hyperparameters or the SABR parameters across expiries. Alternatively, a two-dimensional GP over (K, T) could be used to model $\sigma_0(K, T)$, which would enforce smoothness across expiries.

5 Results

We apply the method described in Section 4 to options on the June 2026 WTI crude oil future (CLM6), using settlement prices from March 13, 2026. We filtered the dataset based on liquidity, choosing strikes with volumes greater than 100 contracts. The dataset consists of 27 strikes ranging from \$56 to \$200 per barrel. We use out-of-the-money options (puts for $K < F_0$, calls for $K \geq F_0$) to compute Black-Scholes implied volatilities via the Black-76 model with $r = 3.7\%$.

To ensure consistency with option prices, we use a forward price inferred from put-call parity, $F_0 = 98.72$, which is consistent across strikes. The time to expiry is $T = 0.17$ years (approximately 63 calendar days to the May 15 options expiry).

5.1 Global SABR Calibration

We fix $\beta = 0.5$ and calibrate σ_0 , α , and ρ by minimizing the sum of squared implied volatility errors across all 27 strikes. The calibrated parameters are:

$$\hat{\sigma}_0 = 2.84, \quad \hat{\alpha} = 1.79, \quad \hat{\rho} = 0.35, \quad \beta = 0.5. \quad (23)$$

Figure 1 shows the globally calibrated SABR smile against the market data. SABR captures the overall shape of the smile well, with a root mean square error (RMSE) of 0.68 volatility points. However, the fit is visibly imperfect. The maximum absolute error is 1.76 volatility points.

5.2 Strike-Dependent σ_0 Inversion

At each observed strike, we invert the SABR formula to find the value of σ_0 that exactly reproduces the market implied volatility, holding $\hat{\alpha}$, β , and $\hat{\rho}$ fixed. Figure 2 plots the resulting function $\sigma_0(K)$.

As anticipated, $\sigma_0(K)$ is nearly flat. Its mean is 2.84 (essentially equal to $\hat{\sigma}_0$), with a standard deviation of only 0.07 and a range of [2.74, 3.05]. This confirms that SABR provides a good structural fit and that the residual correction required from the Gaussian Process is small.

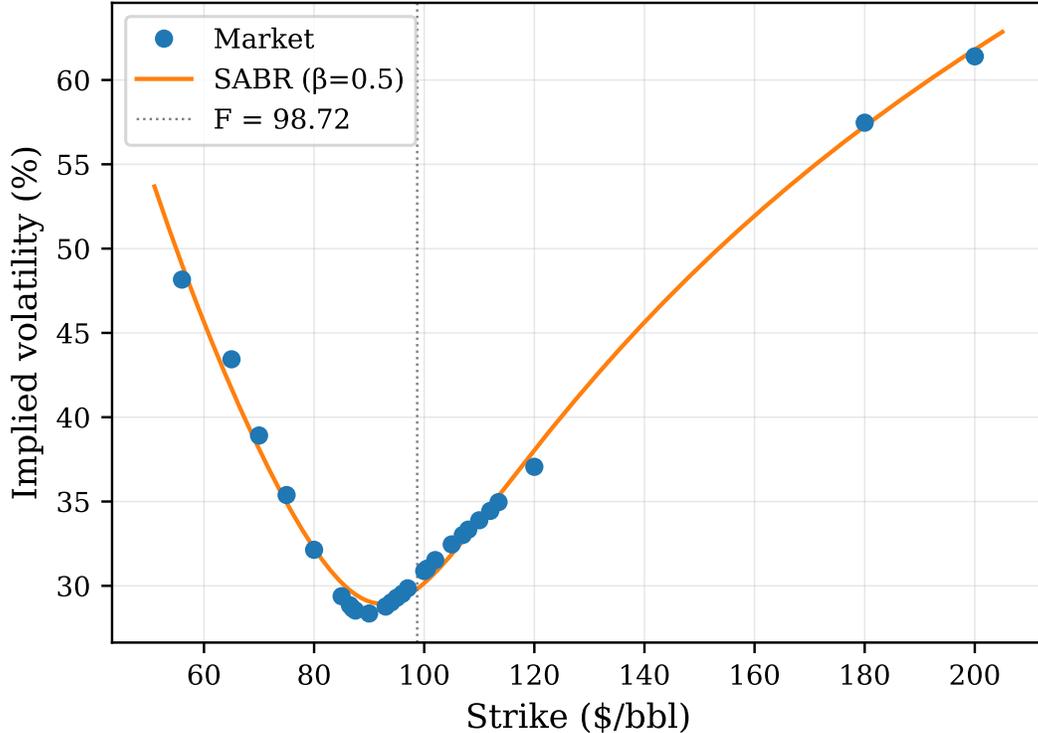


Figure 1: Market implied volatilities (circles) and globally calibrated SABR smile (dashed) for CLM6 options on March 13, 2026. The vertical dotted line marks the forward price $F_0 = 98.72$.

5.3 Gaussian Process Interpolation

We fit a GP with squared exponential kernel to the 27 observations $\{(K_i, \sigma_0^{(i)})\}$, using $\hat{\sigma}_0$ as the constant prior mean. Hyperparameters are estimated by maximizing the log marginal likelihood, which yields a signal standard deviation of $\sigma_f = 0.091$, a length scale of $\ell = 5.4$ (in strike units of \$/bbl), and a noise standard deviation of $\sigma_n = 2.3 \times 10^{-3}$. The very small fitted noise variance confirms that the inversion is effectively exact and the GP is interpolating rather than smoothing.

5.4 Hybrid Smile

Figure 3 shows the resulting hybrid SABR+GP smile, together with the pure SABR fit and the market data. The shaded region represents an approximate $\pm 2\sigma$ confidence band, obtained by propagating the GP posterior standard deviation through the SABR formula via a first-order approximation.

The hybrid model passes through the market data almost exactly, achieving an RMSE of 0.014 volatility points. This is a factor of 49 improvement over the pure SABR fit. The maximum absolute error is 0.031 volatility points. Between observed strikes, the interpolation is smooth, inheriting the smoothness of the GP kernel. In the wings beyond the last observed strikes, the model reverts toward the globally calibrated SABR smile as the GP

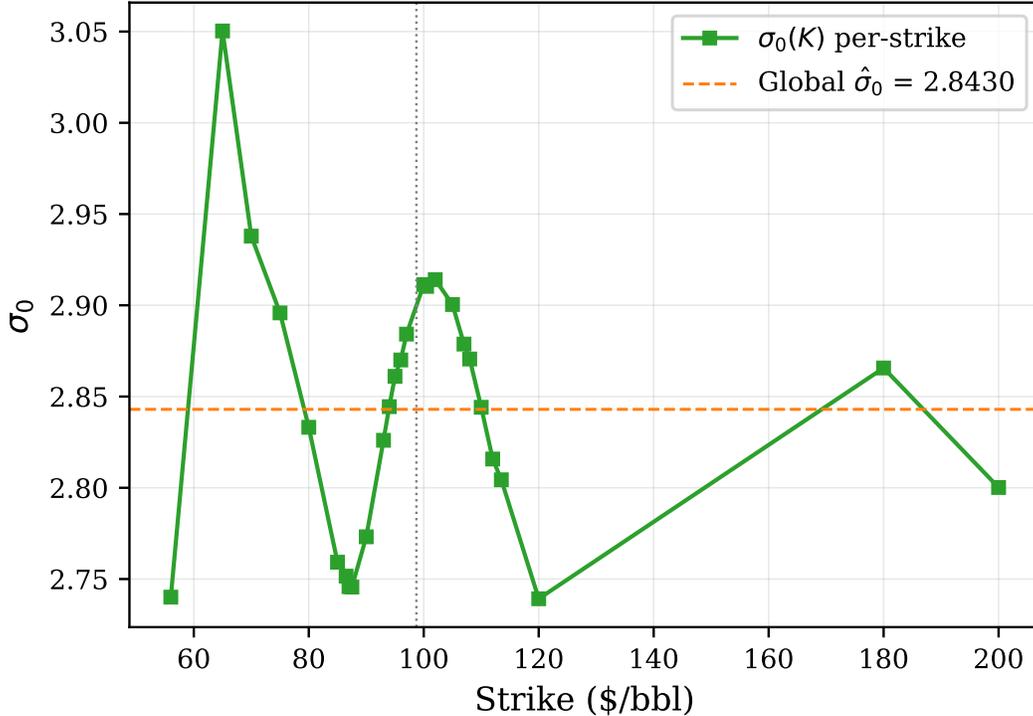


Figure 2: Per-strike values of $\sigma_0(K)$ obtained by inverting the SABR formula at each market strike. The dashed line shows the globally calibrated value $\hat{\sigma}_0 = 2.84$. The near-flatness of this curve is the key property that makes GP interpolation well-behaved.

posterior mean returns to $\hat{\sigma}_0$, providing economically motivated extrapolation.

The confidence band is narrow near observed strikes, where the GP posterior variance is small, and widens in regions without data, providing a natural measure of model uncertainty.

5.5 Fitting Residuals

Figure 4 compares the residuals at observed strikes for the pure SABR fit and the hybrid model. The SABR residuals show systematic structure. They undershoot near the money and overshoot in the wings. This happens because the SABR parametric form cannot fully capture the market smile shape. The hybrid residuals are negligible, confirming that the GP correction successfully absorbs the systematic error.

5.6 Comparison with Direct GP on Implied Volatility

To illustrate the advantage of applying the GP to $\sigma_0(K)$ rather than directly to Black-Scholes implied volatility, Figure 5 compares the two approaches. A GP is fitted directly to the market implied volatilities $\sigma_{\text{mkt}}(K_i)$, using the same squared exponential kernel and marginal likelihood optimization.

The direct GP fits the observed data adequately, but its interpolation and extrapolation behavior is problematic. In regions of sparse data, and beyond the last observed strikes, the

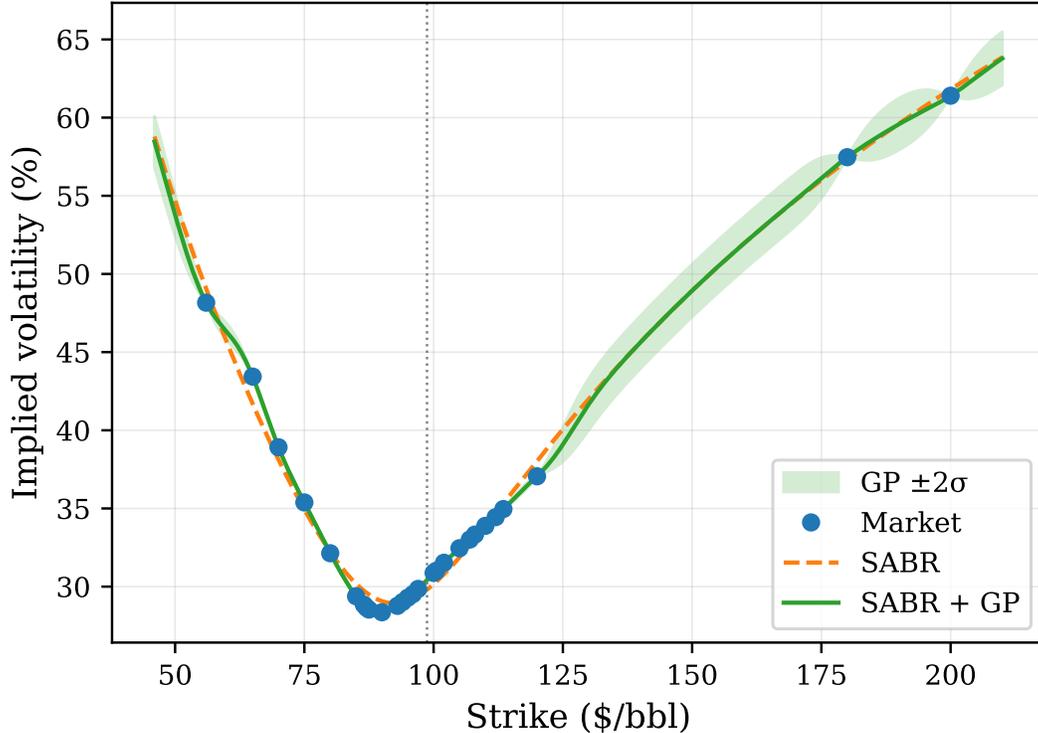


Figure 3: Hybrid SABR+GP implied volatility smile (solid) with approximate $\pm 2\sigma$ confidence band (shaded), compared with pure SABR (dashed) and market data (circles).

predicted volatility reverts to the GP prior mean (approximately the average of the observed volatilities). This produces implausible oscillations in the volatility smile and a downturn in the extrapolated wings. In contrast, the hybrid method’s extrapolation inherits the skew and curvature of the SABR model, which rises in both wings as expected.

6 Conclusions

We have proposed a method for fitting the implied volatility smile that combines the structural backbone of the SABR stochastic volatility model with the interpolation flexibility of Gaussian Process regression. By applying the GP to the SABR initial volatility parameter σ_0 as a function of strike, rather than to Black-Scholes implied volatility directly, we ensure that the GP operates on a nearly flat function, that interpolation is smooth and well-behaved, and that extrapolation into the wings inherits the economically motivated shape of the SABR model.

The method is straightforward to implement given existing SABR and GP libraries, achieves exact calibration at observed market strikes, and provides uncertainty quantification through the GP posterior. We believe it offers a practical improvement over both pure-parametric and pure-nonparametric approaches to smile fitting.

A reference implementation, including the code used to produce the figures in this paper, is available at https://github.com/edparcell/volatility_smile_sabr_gp

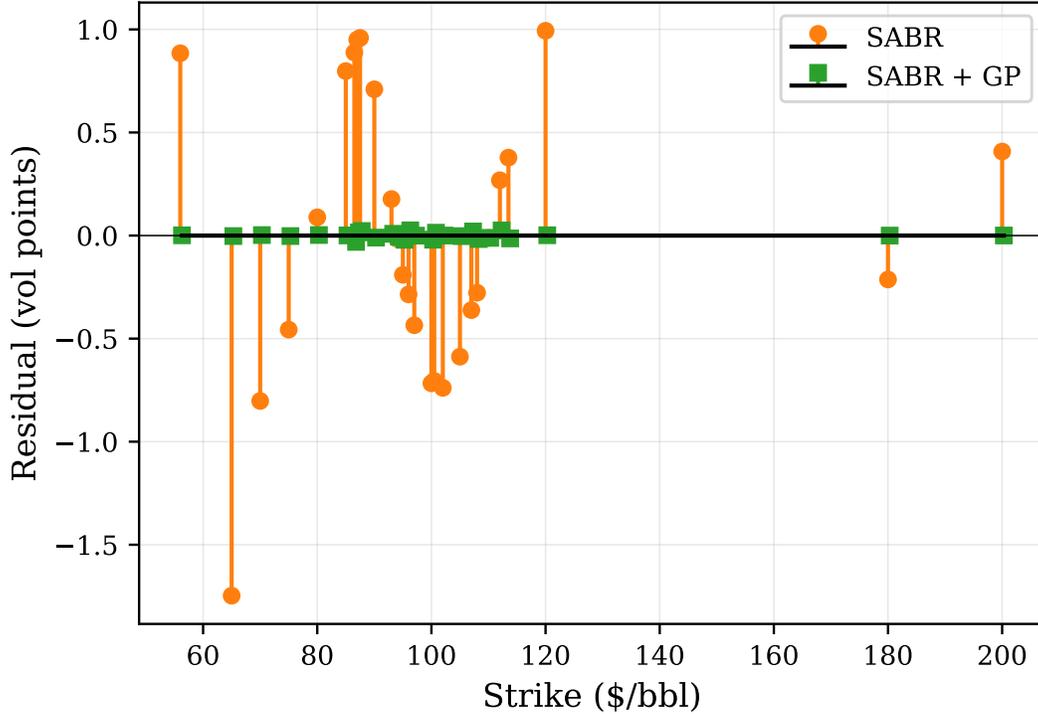


Figure 4: Fitting residuals at market strikes for SABR (circles) and the hybrid SABR+GP model (squares). Note the change in scale: the hybrid residuals are roughly two orders of magnitude smaller.

AI Disclosure

The core methodological idea (applying Gaussian Process regression to the SABR σ_0 parameter rather than to implied volatility directly) was developed by the author circa 2015–2016, and a partial draft was begun in 2020. This paper was completed in 2026 with substantial assistance from Claude (Anthropic), which was used to expand the exposition, implement the method, produce the figures, and survey related literature. The author provided the mathematical framework and reviewed all technical content. The choice to publish with AI assistance, rather than leave the idea unpublished indefinitely, was a deliberate one.

References

- Chataigner, M., Cousin, A., Crépey, S., Dixon, M., and Gueye, D. (2022). Beyond Surrogate Modeling: Learning the Local Volatility Via Shape Constraints. *arXiv preprint arXiv:2212.09957*
- Gatheral, J. (2004). A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. Presentation at Global Derivatives.
- Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. (2002). Managing smile risk. *Wilmott Magazine*, pages 84–108.

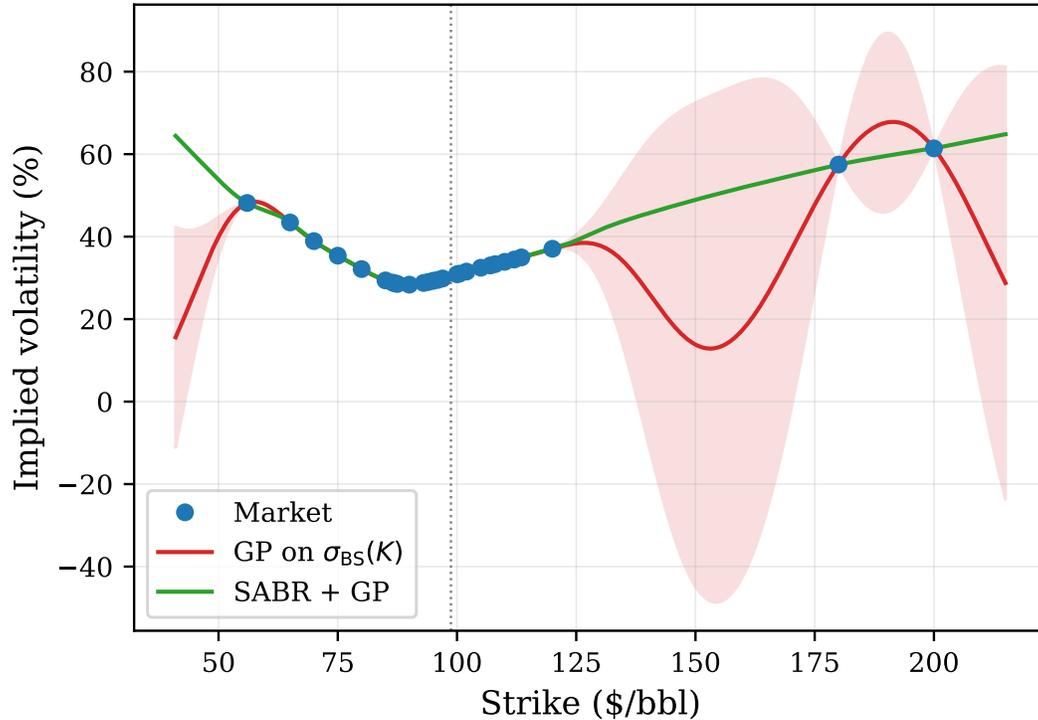


Figure 5: Comparison of the hybrid SABR+GP method with a GP fitted directly to Black-Scholes implied volatilities. The direct GP produces flat extrapolation in the wings, while the hybrid method inherits SABR’s economically motivated wing behavior.

Rasmussen, C. E. and Williams, C. K. I. (2005W). *Gaussian Processes for Machine Learning*. MIT Press.

Zhuang, Y. and Wu, X. (2025). SABR-Informed Multitask Gaussian Process: A Synthetic-to-Real Framework for Implied Volatility Surface Construction. *arXiv preprint arXiv:2506.22888*.

Almeida, C., Fan, J., Freire, G., and Tang, F. (2022). Can machine learning correct option pricing models? Working paper.