

UNIVERSIDADE DE SÃO PAULO  
ESCOLA POLITÉCNICA  
TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO  
DEPARTAMENTO DE ENGENHARIA MECATRÔNICA

Giulia Duo Cardella

**Predição de Superfície de Volatilidade de ativos  
brasileiros utilizando Deep Learning**

São Paulo  
2022

GIULIA DUO CARDELLA

**Predição de Superfície de Volatilidade de ativos  
brasileiros utilizando Deep Learning**

— Versão Original —

Monografia apresentada como Trabalho de  
Conclusão de Curso para a Graduação em  
Engenharia Mecatrônica da Escola Politéc-  
nica da Universidade de São Paulo.

Orientador: Profa. Dra. Larissa Driemeier

São Paulo  
2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

#### Catalogação-na-publicação

Cardella, Giulia Duo

Predição de Superfície de Volatilidade de ativos brasileiros utilizando Deep Learning/ G.D.Cardella – São Paulo, 2022.  
150p.

Monografia (Graduação em Engenharia Mecatrônica) – Escola Politécnica da Universidade de São Paulo. Graduação em Engenharia Mecatrônica.

1. volatilidade 2. deep learning 3. SVI 4. B3 5. Ações 6. Superfície de Volatilidade Implícita.  
I. Universidade de São Paulo. Escola Politécnica. Graduação em Engenharia Mecatrônica II.t.

Para meus pais, Julio e Claire, minha irmã Giovanna e meu companheiro Vítor.

# Agradecimentos

Agradeço à Prof. Dr<sup>a</sup> Larissa Driemeier, pelos ensinamentos.

When you walk through a storm  
Hold your head up high  
And don't be afraid of the dark  
At the end of a storm  
There's a golden sky  
And the sweet silver song of a lark  
Walk on through the wind  
Walk on through the rain  
Though your dreams be tossed and blown.

— *Gerry and the Pacemakers*

# Sumário

**Sumário** • *v*

**Resumo** • *vii*

**Abstract** • *viii*

**Lista de Figuras** • *ix*

**Lista de Tabelas** • *x*

**1 Introdução** • *1*

**2 Revisão Bibliográfica** • *4*

**3 Métodos** • *8*

3.1 Black-Scholes Merton Inverso • *8*

3.1.1 Método de Brent • *9*

3.2 Parametrização da IVS • *9*

3.2.1 Otimização por Evolução Diferencial • *11*

3.3 ConvLSTM • *13*

**4 Dados** • *15*

4.1 Obtenção dos Dados • *15*

4.2 Tratamento dos Dados • *16*

4.3 Obtenção da SELIC • *16*

4.4 Moneyness • *17*

4.5 Volatilidade Implícita • *18*

4.6 Variância • *18*

**5 Modelo** • *20*

5.1 Requisitos • *20*

5.2 Dinâmica da IVS • *21*

5.2.1 Interpolação • *21*

5.3 Parâmetros da Rede • *23*

5.3.1 ConvLSTM • *23*

5.3.2 Inputs e Outputs • *23*

5.4 Modelo Benchmark - Naive Model • *24*

5.5 Cálculo de Erro • *25*

**6      Resultados • 26**

6.1    Comportamento dos Fatores de Risco • 26

6.2    Predição dos Fatores de Risco • 30

6.3    Análise dos Resultados • 33

    6.3.1 Inerentes à realidade brasileira • 34

    6.3.2 Ativos subjacentes diferentes • 37

**7      Conclusão • 39**

**Referências • 41**

**Apêndices • 44**

**Anexos • 45**

## Resumo

CARDELLA, G.D. *Predição de Superfície de Volatilidade de ativos brasileiros utilizando Deep Learning*. 2022. Monografia (Graduação em Engenharia Mecatrônica) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Mecatrônica., 2022.

O principal objetivo desta monografia é prever a Superfície de Volatilidade Implícita (*IVS*) de ativos brasileiros, sendo pioneiro na aplicação de *Deep Learning* na análise de ações listadas na B3 (Bolsa do Brasil). Assim, foram utilizados dados históricos de cotações de opções de PETR4 para obtenção das Volatilidades Implícitas através do método de Black and Scholes Inverso. As volatilidades foram, em seguida, *ajustadas* em um modelo paramétrico e estocástico denominado *SVI*, cujos parâmetros correspondem a fatores de risco. Por fim, os parâmetros foram utilizados no treinamento com múltiplas *ConvLSTM*, em um modelo final capaz de capturar a correlação espaço-temporal entre *strikes* e *tenors*. Os resultados foram analisados através de métricas de erro (*MAE*), além da comparação com um modelo *benchmark* simples. O método mostrou bons resultados para opções com vencimentos de curto prazo e para previsões inferiores a 10 dias, já para outros *tenors* e previsões, os resultados podem ser afetados por fatores inerentes a conjuntura brasileira. Portanto, conclui-se que, apesar das especificidades e limitações, é possível utilizar técnicas de Inteligência Artificial para previsão da *IVS* de ativos listados na B3.

**Palavras-chave:** Volatilidade. Deep Learning. *SVI*. B3. Ações. Superfície de Volatilidade Implícita.

## Abstract

CARDELLA, G.D. *Predicting Volatility Surface of Brazilian Assets using Deep Learning*. 2022. Monografia (Graduação em Engenharia Mecatrônica) – Escola Politécnica da Universidade de São Paulo. PECE – Programa de Educação Continuada em Engenharia. University of São Paulo, São Paulo, Brazil. 2022.

The main purpose of this monograph is to predict the Implied Volatility Surface (*IVS*) of Brazilian assets, being a pioneer in the application of *Deep Learning* in the analysis of stocks listed on B3 (Brazilian Exchange). Thus, historical data of PETR4 options were used to obtain the Implied Volatilities through the Inverse Black and Scholes method. The volatilities were then *fitted* in a parametric and stochastic model called *SVI*, whose parameters correspond to risk factors. Finally, the parameters were used to train multiple *ConvLSTM* layers, resulting in a final model capable of capturing the space-time correlation between *strikes* and *tenors*. The results were analyzed using error metrics (*MAE*), in addition to a comparison with a simple *benchmark* model. The method was successful for short-term options and for predictions within 10 days, whilst other *tenors* and predictions may be affected by factors inherent to the Brazilian scenario. Therefore, it is concluded that, despite the specificities and limitations, it is possible to use Artificial Intelligence techniques to predict the *IVS* of B3 listed stocks.

**Keywords:** Volatility. Deep Learning. SVI. B3. Stocks. Implied Volatility Surface.

# Listas de Figuras

1.1	Sorriso de Volatilidade. <i>Fonte: Própria.</i>	2
1.2	Superfície de Volatilidade. <i>Fonte (Cont e Paris 2002)</i>	3
3.1	Volatilidade obtida através de <i>BSMI</i>	9
3.2	Volatilidadade Histórica de PETR3 e PETR4 (2022). <i>Fonte: InvestMax</i>	9
3.3	Superfícies <i>fittadas</i> x Dados Experimentais para 20/08/2015. <i>Fonte: Própria.</i>	12
3.4	Recurrent Neural Network <i>Fonte: Mishra, Agarwal e Puri 2018</i>	13
3.5	ConvLSTM <i>Fonte: Stollenga 2016</i>	14
4.1	Arquivo de Dados Históricos B3 (2017).	15
4.2	SELIC - Taxa de Juros Básica - (2010-2022). <i>Fonte: Própria.</i>	17
4.3	<i>Log-Forward Moneyness.</i> <i>Fonte: Própria</i>	17
4.4	Superfície de Volatilidade - PETR4. <i>Fonte: Própria.</i>	18
4.5	Variâncias Totais - PETR4. <i>Fonte: Própria.</i>	19
5.1	Exemplo de parâmetro (a) ao longo do tempo para <i>tenor = 7</i> . <i>Fonte: Própria.</i>	22
5.2	Sorriso de Volatilidade Interpolado. <i>Fonte: Própria.</i>	22
5.3	Sorriso de Volatilidade Tenors Fixos. <i>Fonte: Própria.</i>	23
6.1	Exemplo de parâmetros para 04/01/2010 <i>Fonte: Própria.</i>	26
6.2	Comportamento dos fatores de risco <i>Fonte: Própria.</i>	27
6.3	Comportamento de a ao longo do tempo <i>Fonte: Própria.</i>	28
6.4	Comportamento de b ao longo do tempo <i>Fonte: Própria.</i>	29
6.5	Comportamento de $\sigma$ ao longo do tempo <i>Fonte: Própria.</i>	29
6.6	Comportamento de m ao longo do tempo <i>Fonte: Própria.</i>	30
6.7	Comportamento de $\rho$ ao longo do tempo <i>Fonte: Própria.</i>	30
6.8	Parâmetros de treinamento. <i>Fonte: Própria.</i>	31
6.9	Convergência das <i>MAEs.</i> <i>Fonte: Própria.</i>	32
6.10	Predição de valores <i>at-the-money.</i> <i>Fonte: Própria.</i>	33
6.11	Exemplo de Sorriso previsto vs Realizado e <i>Naive</i> para o dia 07/05/2018, prevendo 5 dias à frente. <i>Fonte: Própria.</i>	33
6.12	SELIC vs Fed Funds. <i>Fonte: Própria.</i>	35
6.13	Número de <i>strikes</i> por <i>tenor.</i> <i>Fonte: Própria.</i>	36
6.14	Sorriso de Volatilidade de BOVESPA. <i>Fonte: Yamamoto (2017).</i>	37
6.15	Efeito de notícias. <i>Fonte: Valor e InvestNews.</i>	38
A.1	Estrutura de Arquivos. <i>Fonte: Própria.</i>	44

# Listas de Tabelas

2.1	Resultados por Autor e Modelo . . . . .	6
4.1	Convenção de Vencimentos <i>Fonte: B3 Expiry Agenda</i> . . . . .	16
6.1	MAE no dataset de treino (05-01-2010, 31-05-2016) e de teste (01-06-2016, 09-09-2022) . . . . .	32
6.2	MAE Relativa (Model - Naive) no dataset de treino (05-01-2010, 31-05-2016) e de teste (01-06-2016, 09-09-2022) . . . . .	32

# Introdução

Opções são contratos financeiros nos quais o titular tem o direito, mas não a obrigação de, na data de vencimento, comprar (*call*) ou vender (*put*) um determinado ativo. As três principais características de uma opção são: o *strike* (preço de exercício), o ativo objeto e a data de vencimento (também conhecido como *tenor*).

Para que opções sejam negociadas, é necessário especificá-las. Com isso, diversos modelos foram criados, como Black-Scholes Merton (*BSM*) (Black e Scholes 1973), Heston, simulação de Monte Carlo e outros. Dentre os citados, o modelo mais utilizado para especificação de opções é o *BSM*, definido pela equação 1.1, de acordo com Hull (2009). Ela é obtida através de uma equação diferencial parcial complexa, cuja derivação não é parte do escopo desta monografia.

Apesar de diferentes, todos os modelos utilizados na definição do valor de uma opção dependem de um componente: a volatilidade do ativo, que pode ser definida como a variação no retorno do ativo em questão. Existem dois tipos de volatilidade: a histórica e a implícita. A volatilidade implícita (*IV*) pode ser descrita como a expectativa de mercado para a volatilidade. Já a histórica (*RV*) é definida como a volatilidade realizada em um determinado período.

$$C(S, t) = N(d_1)S - N(d_2)Ke^{-rt} \quad (1.1)$$

$$d_1 = \frac{1}{\sigma\sqrt{t}} \left[ \ln\left(\frac{S}{K}\right) + t \left( r + \frac{\sigma^2}{2} \right) \right] \quad (1.2)$$

$$d_2 = \frac{1}{\sigma\sqrt{t}} \left[ \ln\left(\frac{S}{K}\right) + t \left( r - \frac{\sigma^2}{2} \right) \right] \quad (1.3)$$

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}z^2} dz \quad (1.4)$$

Nas equações 1.1 a 1.4,  $C$  é o preço de uma opção *Call*,  $S$  é o preço atual do ativo (preço de mercado),  $K$  é o *strike*,  $r$  é a taxa de juros livre de risco (*risk-free interest*), que, no Brasil, corresponde à *SELIC*,  $\sigma$  é a volatilidade implícita do ativo,  $t$  o tempo até a maturidade/vencimento e  $N$  é uma função de distribuição normal. Há uma derivação semelhante para o preço de opções *Put*, que pode ser encontrada no Apêndice [1].

Cada um dos modelos de precificação tem suas hipóteses acerca da volatilidade de um ativo. *BSM*, por exemplo, assume que a volatilidade é constante e que todos *strikes* possuem o mesmo nível.

Porém, *traders* (negociantes) de opções e pesquisadores notaram que, ao acompanhar e obter a volatilidade dos ativos através dos preços negociados das opções, a *IV* variava tanto com relação aos *Strikes* quanto aos vencimentos.

Com isso, Rubinstein 1994 e Jackwerth e Rubinstein 1996 iniciaram pesquisas com respeito à distribuição de *IV* em opções sobre ações e constataram que a *IV* apresentava uma distorção ao comparar diferentes *strikes* (Hull, Suo e Daglish 2001). Essa distorção é nomeada Sorriso de Volatilidade (*Volatility Smile*) e é representado como um gráfico entre Volatilidade x *Strikes* (Medvedev 2019). Um exemplo de Sorriso de Volatilidade pode ser encontrado na Figura 1.1. Um sorriso de volatilidade é composto pelo *skew* de volatilidade, pela *call wing* e pela *put wing*, que são as extremidades do Sorriso.

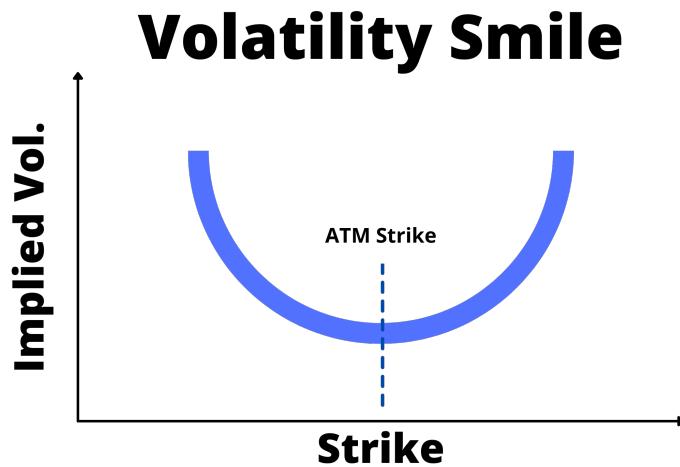


Figura 1.1: Sorriso de Volatilidade. *Fonte: Própria.*

Ainda em contraste com as hipóteses do *BSM*, que assume que as volatilidades são constantes entre vencimentos, de acordo com Dupire 1994 e Dumas, Fleming e Whaley 1996, a volatilidade varia também com a maturidade, aumentando ou decaendo até a data de vencimento (Hull, Suo e Daglish 2001). Essa propriedade da volatilidade implícita é conhecida como Estrutura de Termo da Volatilidade (*Volatility Term Structure*).

Ao combinar o Sorriso com a Estrutura de Termo da Volatilidade, é possível obter uma figura 3D, conhecida como superfície de volatilidade (IVS). Um exemplo de IVS pode ser encontrado na Figura 1.2.

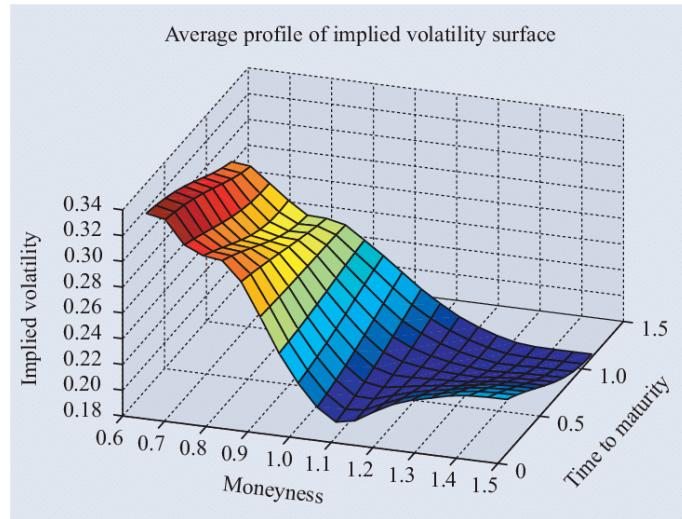


Figura 1.2: Superfície de Volatilidade. *Fonte (Cont e Paris 2002)*

A IVS é utilizada no mercado de derivativos de ações para diversas funções, dentre as quais estão incluídas análise e gestão de risco, criação de estratégias, proteção de posições e também para entender a expectativa futura que os constituintes do mercado têm acerca de um determinado ativo (Ahoniemi 2006). É um assunto que tem sido estudado com profundidade por bancos de investimento, gestoras de capitais, fundos de investimento e outras áreas dentro do Mercado Financeiro. Prever a IVS de um ativo pode ser uma vantagem competitiva para os participantes do mercado, que serão capazes de antecipar movimentos, além de ter bases para tomada de decisões.

Para prever e antecipar a variação da IV, diversos modelos foram utilizados. Como por exemplo *GARCH* (*Generative Autoregressive Conditional Heteroskedasticity*) e o *EWMA* (*Exponentially Weighted Moving Average*), ambos modelos estatísticos utilizados para previsão de séries temporais. Apesar de poderosos, são modelos capazes de prever apenas a volatilidade do strike chamado *at-the-money* (o strike  $K$  mais próximo aos *Spots S*). Ao prever apenas um valor pontual do Sorriso de Volatilidade, parâmetros relevantes como o *skew* são perdidos (Medvedev 2019). Com isso, para evitar perda de informações importantes para o mercado, técnicas de *Machine Learning* e Inteligência Artificial começaram a ser utilizadas para previsão da IVS.

O principal objetivo desta monografia é prever a Superfície de Volatilidade Implícita (IVS) de ativos brasileiros do mercado de ações.

## Revisão Bibliográfica

Conforme o conceito de *IV* ganhou importância no Mercado Financeiro, diversos estudos visando modelar e prever a *IVS* através dos *strikes* e dos vencimentos foram desenvolvidos, em especial para opções do Índice *S&P*, um dos ativos mais negociados. Porém, prever a *IVS* de ativos não é uma tarefa trivial.

Apesar da *IVS* ser caracterizada como a expectativa de retorno futuro de uma ação, os preços das opções não representam fielmente uma antecipação dos retornos futuros. Isso é dado pelo fato da variação da *IV* ser estocástica (Dumas, Fleming e Whaley [1996](#)). Por isso, a previsão da *IVS* é uma tarefa complexa (Medvedev [2019](#)).

Grande parte da complexidade do problema se dá pela sua alta dimensionalidade. Para se estimar o comportamento futuro (predição dos fatores de curto prazo) da superfície de volatilidade, uma possível solução é a redução de dimensionalidade, projetando a superfície em fatores de risco e assumindo dinâmicas para estes. Esses fatores de risco estocásticos formam um vetor de base para a superfície de volatilidade implícita (*IVS*). Exemplos na literatura incluem os primeiros três autovetores de uma Análise de Componentes Principais (*PCA*), os parâmetros de um polinômio, ou, quaisquer outros parâmetros de modelos de volatilidade estocásticos. Assim, a dinâmica da *IVS* é dada pelos fatores de risco de volatilidade, que podem ser estimados por modelos estatísticos ou pela premissa de processos estocásticos.

Na literatura, em se tratando de modelos estatísticos e métodos matemáticos, diversos estudos foram feitos. Como exemplo há Dumas, Fleming e Whaley ([1996](#)) que apresentaram algoritmo de Árvore de Decisão para prever as volatilidades do Índice *S&P*. Os resultados obtidos apresentaram um *AVE* (*Average Error*) de -0,010 e um *RMSE* (*Root Mean Squared Error*) de 0,227. Audrino e Colangelo ([2011](#)) propuseram um método semi-paramétrico, combinando um modelo baseado em Dumas, Fleming e Whaley ([1996](#)) com um algoritmo de *boost* (*fitting* de árvores de regressão) para melhorar os resultados, o que reduziu o erro do modelo inicial em 31%. Posteriormente, Chen, Zhou e Li ([2016](#))

propuseram um modelo dinâmico que era capaz de descrever o comportamento da *IVS* ao longo do tempo. Para o estudo, os autores usaram Filtro de Kalman para propagar, através do tempo, um sistema não linear, composto pela função paramétrica de uma *IVS* e seus componentes.

Em outras Bolsas de Ações, estudos similares surgiram. Wang et al. (2017) apresentaram um modelo de fator dinâmico restrito para prever a *IVS* do *ETF SSE 50 (Shanghai StockExchange 50ETF)*. Os autores modelaram as séries temporais dos fatores através de um processo estocástico chamado *Ornstein-Uhlenbeck (OU)*. Parâmetros foram estimados usando um Filtro de Kalman, os quais mostraram uma performance acima da expectativa quando comparados a modelos tradicionais de *IVS*.

Para opções de ações do tipo americanas, Bernales e Guidolin (2014) usaram um método de dois passos para construir uma *IVS*. No primeiro passo, os autores caracterizaram as *IVSs* de opções de ações através do *fitting* em um modelo de volatilidade determinístico. Posteriormente foi estimado um modelo *VARX (Vector Autoregressive with Exogenous Variable)* para prever a *IVS* de opções de ações. O modelo implementado no segundo passo é construído usando o histórico de coeficientes do modelo determinístico usado no primeiro.

Em se tratando de previsões da dinâmica da *IVS* com *Machine Learning (ML)*, estudos recentes, a partir de 2018, têm sido pioneiros. Por exemplo, Sun (2018), que testou diferentes modelos de *ML* para prever as mudanças na *IVS* de opções do Índice S&P num horizonte de um dia até um mês.

Já Huynh (2018) sugeriu que uma rede neural de múltiplas camadas (*multi-layer perception neural network*) superaria modelos autorregressivos (tal qual o *VARX*) e *Random Walk*, estudando a dinâmica da volatilidade do Índice *Swedish OMXS 30*. O autor concluiu que era mais difícil prever opções dentro do dinheiro (*in-the-money* ou *ITM*) do que fora do dinheiro (*out-of-the-money* ou *OTM*), e que prever opções de curto prazo era uma tarefa mais complicada do que prever opções de longo prazo.

Chen e Zhang (2020) introduziu um *attention mechanism* em um modelo *LSTM* para prever o *IVS* um passo à frente. *LSTM* é uma rede neural recorrente, que se mostrou efetiva na previsão de *datasets* temporais e é comumente utilizada para prever preços de ações. No modelo do autor, a previsão é feita nos pontos discretos do *IV Smile* (numa discretização de 45 pontos) com o *Moneyness* (percentual do strike em relação ao preço atual do ativo) entre 80% e 120% e vencimentos em Março, Junho, Dezembro. Em todo tempo *t*, as variáveis explicativas são *IVs* observadas somadas às médias semanais e mensais de cada um dos 45 pontos. Para demonstrar convergência, os autores usaram o erro quadrático médio (*MAE*) no modelo de teste.

Medvedev (2019) modelou toda a *IVS* usando arquiteturas de redes neurais recorrentes. Ele propôs um modelo de rede *LSTM* convolucional (*ConvLSTM*) para produzir previsões multivariadas e de múltiplos passos para a *IVS* do Índice *S&P 500* com horizontes de 1 dia, 30 dias e 90 dias.

Autor	Modelo	RMSE ou MAE	Predição
Medvedev (2019)	ConvLSTM	0.0156	IVS
Bloch e Book (2021)	SVI + ConvLSTM	0.004584*	IVS
Wang et al. (2017)	State Space Model	0.0713	IVS
Wang et al. (2017)	Random Walk Model	0.0813	IVS
Zhu (2018)	Garch(1,1)	0.05244	IV value
Zhu (2018)	EMWA(1,1)	0.05668	IV value
Zhu (2018)	Garch&EMWA	0.05194	IV value
Huynh (2018)	Neural Network	0.01839	IVS
Huynh (2018)	Vector Autoregressive	0.02395	IVS

\*Diferença entre o modelo proposto e um modelo *naive* usado como *benchmark*

Tabela 2.1: Resultados por Autor e Modelo

Ainda com relação à predição de *IVS* utilizando conceitos de *Machine Learning*, *Deep Learning* e Inteligência Artificial, Bloch e Book 2021 propõem que para tratar os problemas supracitados deve-se evoluir a *IVS* em múltiplos passos à frente para todos os *strikes* e vencimentos. Para lidar com a alta dimensionalidade do problema, projeta-se a superfície em algumas funções de base, resultando numa *IVS* de *grid* fixo. A fim de simplificar ainda mais a modelagem, os autores utilizam a parametrização do Sorriso de Volatilidade pelo modelo *SVI* e deixam cada parâmetro em forma de estrutura de termos. Esses parâmetros são montados em dados históricos para fixar os *strikes* e vencimentos. São exploradas diferentes arquiteturas de *Deep Recurrent Networks* a fim de representar as variáveis explicativas num vetor contexto que será o *input* de uma *Deep Network*. Entretanto, como as relações espaço-temporais da *IVS* são perdidas, múltiplas camadas de *ConvLSTM* são somadas para formar uma estrutura de codificação e previsão, culminando em uma rede capaz de compreender as relações espaço-temporais entre *strikes* e vencimentos.

Após a previsão, na proposta de Bloch e Book 2021, a *IVS* é reconstruída através do Sorriso paramétrico, onde cada parâmetro é substituído pelo seu valor previsto. Então, a modelagem paramétrica da *SVI* é usada para inferir a volatilidade para determinados *strikes* e vencimentos, possibilitando recuperar o preço de mercado através da fórmula de *BSM*. A performance do modelo é explorada em comparação a uma estratégia simples (*Regressão de Naive*) e computando uma medida de precisão (*MAE* e *RMSE*) para o Índice

*S&P 500*. Nos casos de previsões de longo prazo para vencimentos curtos e médios a expectativa foi superada.

Nesta monografia, o método a ser utilizado consiste em replicar o algoritmo proposto por Bloch e Book [2021](#) ao cenário brasileiro. Assim, as seguintes etapas são propostas:

- Obtenção das Volatilidades Implícitas em preços de mercado de opções através do método de *Black and Scholes Inverso*;
- Obtenção de Fatores de Risco paramétricos do modelo *SVI*;
- Obtenção de um Grid Fixo de Parâmetros;
- Predição dos Fatores de Risco utilizando *ConvLSTM*;
- Análise dos resultados através de *MAE*.

Quanto à escolha da rede, no trabalho de Bloch e Book [\(2021\)](#), a *ConvLSTM* provou ser um modelo eficiente para prever *IVS* através do tempo. O modelo apresentou o menor *MAE* e comparado aos outros, o que pode ser visto em [2.1](#).

Até o momento, nenhuma monografia tratando a aplicação de Inteligência Artificial (IA) na previsão de *IVSs* de ativos brasileiros foram encontrados, tornando este trabalho um pioneiro para o cenário.

# CAPÍTULO 3

---

## Métodos

### 3.1 Black-Scholes Merton Inverso

Como mencionado na seção 1, a fim de especificar uma opção, são necessários 5 *inputs*, sendo eles:  $S$ ,  $K$ ,  $r$ ,  $\sigma$  (volatilidade implícita) e  $t$ . Porém, para predição de IVS, o objeto de estudo não são os preços de opções e, sim, as volatilidades. Com isso, é necessário capturar, através dos preços de mercado, as volatilidades implícitas em cada uma das opções.

Para isso, é necessário resolver um problema inverso: o *BSM Inverso*. O problema a ser resolvido para encontrar a volatilidade implícita nos preços de opções negociados consiste em um problema de minimização de erros. A ideia principal é minimizar a diferença absoluta entre o preço de mercado e o preço teórico calculado através da equação de *BSM* (Black e Scholes 1973). Essa minimização é feita através do método híbrido, descrito na seção 3.1.1, iterando diferentes valores para a volatilidade implícita dentro de um intervalo definido: [0.01, 6]. O intervalo foi escolhido de acordo com volatilidades históricas (*RV*) dos ativos, um exemplo de *RV* para Petrobras está na Figura 3.2.

A minimização a ser feita pode ser encontrada na equação 3.1:

$$f(x) = 0 \quad (3.1)$$

Em que

$$f(x) = |C_{market} - C_{BS}| \quad (3.2)$$

No repositório indicado no Apêndice 7, há a implementação para solução do problema inverso.

	Ticker	Spot	SELIC	Log-Forward	Moneyness	Volatilidade	Dias até vencimento
0	PETRA1	9.57	0.1175		0.103016	0.654427	11
1	PETRA10	9.57	0.1175		-0.110558	0.726080	11
2	PETRA12	9.57	0.1175		0.294520	0.748137	11
3	PETRA13	9.57	0.1175		0.369125	0.809352	11
4	PETRA14	9.57	0.1175		0.192682	0.678424	11
5	PETRA15	9.57	0.1175		0.438550	0.934966	11
6	PETRA18	9.57	0.1175		0.621877	0.999994	11
7	PETRA2	9.57	0.1175		-0.672209	0.999994	11
8	PETRA23	9.57	0.1175		-0.043193	0.681915	11
9	PETRA25	9.57	0.1175		-0.000678	0.646949	11

Figura 3.1: Volatilidade obtida através de *BSMI*

. *Fonte: própria.*

♦ Ações	♦ Cotação	▲ 1 ano	♦ 6 meses	♦ 3 meses	♦ 32 dias	♦ 1 mes	♦ 10 Dias
PETR3	28,12	44,57%	50,98%	51,33%	56,59%	59,74%	63,90%
PETR4	24,25	46,32%	53,53%	54,38%	62,02%	66,26%	75,20%
média		45,45%	52,26%	52,86%	59,31%	63,00%	69,55%

Figura 3.2: Volatilidadade Histórica de PETR3 e PETR4 (2022). *Fonte: InvestMax*

### 3.1.1 Método de Brent

O método de Brent consiste em uma abordagem híbrida proposta por Brent (2013) para aproximar  $f(x)=0$  que inclui elementos dos métodos da secante e da bisseção. É uma técnica de "enclosure", como a bissecção, que começa com um intervalo inicial onde  $f$  muda de sinal e, com o passar das iterações, estabelece uma série de intervalos que compartilham esse aspecto e diminuem de tamanho.

## 3.2 Parametrização da IVS

Dentre as possibilidades de modelos paramétricos para a IVS, será utilizado neste trabalho o movelo SVI (*Stochastic Volatility Inspired*). A escolha de modelo se deu pela possibilidade do SVI de estabelecer condições explícitas nos seus parâmetros, tornando possível a criação de um modelo de preços livre de arbitragens. Além disso, é possível utilizar dados reais de mercado para calibrar a superfície (Aurell 2014).

Como mencionado, o SVI é um modelo paramétrico e estocástico. Um modelo estocástico é definido como indeterminado e gerado por eventos aleatórios.

No modelo *SVI raw*, a variância total escalada no tempo é representada como:

$$\Xi_{SVI}^2(x, t) = a + b(\rho(x - m) + \sqrt{(x - m)^2 + \sigma^2}) \quad (3.3)$$

Em que  $x$  é a *log-forward moneyness*:

$$x = \ln\left(\frac{K_t}{F(t, T)}\right) \quad (3.4)$$

E  $F$  representa o valor do spot atual levado a valor futuro (à data de vencimento em questão) utilizando a taxa de juros livre de risco (*SELIC*).

Além disso:

$$\omega(K, T) = (T - t)\Sigma^2(K, T) \quad (3.5)$$

E:

$$\Xi(K, T) = \sqrt{\omega(\eta, T)} \quad (3.6)$$

O *SVI raw* é a versão mais utilizada de parametrização do *SVI*, além de ser a parametrização original proposta. Os parâmetros da equação 3.3 representam:

- $a$  corresponde à translação vertical no Sorriso.
- $b$  corresponde às inclinações das *wings* de *put* e *call* (extremidades à esquerda e à direita do Sorriso, respectivamente).
- $\rho$ , também chamado de correlação, representa a rotação do Sorriso em torno de seu ponto central (*at-the-money strike*).
- $m$  corresponde à translação horizontal do Sorriso.
- $\sigma$  representa a curvatura do Sorriso, ou seja, um aumento do parâmetro representa uma redução na curvatura *at-the-money* do Sorriso.

Para que não ocorram arbitragens de calendário ou de *butterfly* (ver Anexo), condições para cada um dos parâmetros do modelo são propostas (Aurell 2014):

$$a \geq 0 \quad (3.7)$$

$$0 \leq b \leq \frac{4}{\tau * (1 + |\rho|)} \quad (3.8)$$

$$-1 \leq \rho \leq 1 \quad (3.9)$$

$$\min(x_i) \leq m \leq \max(x_i) \quad (3.10)$$

$$0 < \sigma \leq 10 \quad (3.11)$$

### 3.2.1 Otimização por Evolução Diferencial

*Differential Evolution*, em português Otimização por Evolução Diferencial é um método heurístico capaz de resolver problemas contínuos de otimização que foi proposto por Storn e Price (1997).

A Evolução Diferencial é um algoritmo simples, mas com alto poder, utilizado para otimização numérica. De acordo Cheng e Hwang (2001), tem como principais características e benefícios:

- O algoritmo é de fácil compreensão;
- É eficaz para solucionar problemas de otimização cuja função a ser otimizada não é contínua;
- É um algoritmo que realiza busca estocástica;
- Se mostra eficaz para lidar com pequenas populações;
- Utiliza diferentes regiões como espaço de busca.

A otimização por Evolução Diferencial neste trabalho foi utilizada para obtenção dos fatores de risco da IVS, a partir dos dados históricos de preços e volatilidades, cujos métodos de obtenção se encontram na seção 3.1.

Na Figura 3.3 se encontram três exemplos de Sorrisos de Volatilidade otimizados através do método supracitado para Petrobrás no ano de 2015.

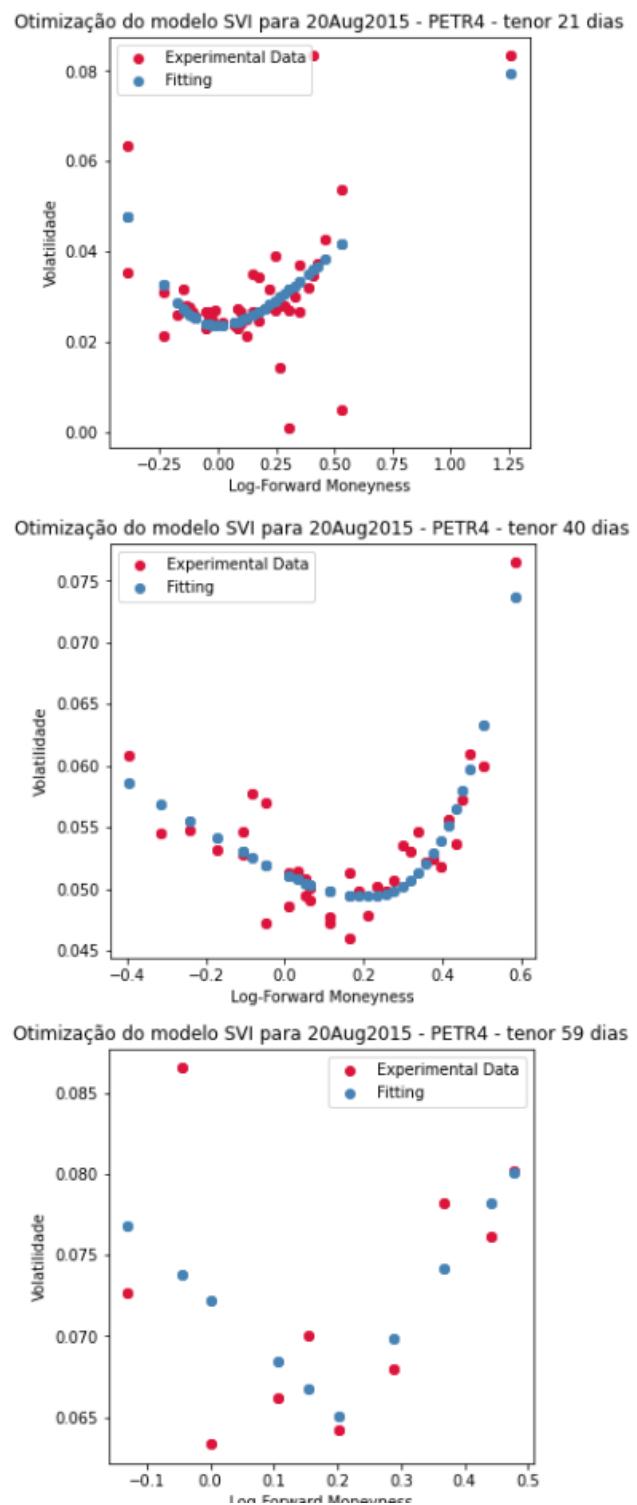


Figura 3.3: Superfícies *fittadas* x Dados Experimentais para 20/08/2015. *Fonte: Própria.*

### 3.3 ConvLSTM

Redes Neurais Recorrentes (*Recurrent Neural Networks ou RNNs*) são redes capazes de lidar com dados de séries temporais que apresentam correlação com o tempo. As *RNNs* (como *GRU* e *LSTM*) são úteis para prever dados em sequência: elas podem processar os dados em tempo real, um passo por vez, para prever o próximo. As redes aprendem como estimar a distribuição de probabilidade do próximo item da sequência a partir dos dados anteriores. A arquitetura de uma *RNN* é mostrada na Figura 3.4.

Uma rede neural convolucional (*Convolutional Neural Network ou CNN*) é uma rede em que os dados de *input* (usualmente imagens) atravessam camadas de convolução, onde informações importantes serão obtidas por filtros. Shi et al. 2015 foram os primeiros a combinar *CNNs* e *RNNs*, eles usaram a *ConvLSTM* (combinação *LSTM* com estruturas convolucionais) para o *nowcasting* de precipitação. Redes *ConvLSTM* são comumente usadas para previsões espaço-temporais e um exemplo do tipo de arquitetura pode ser visto na figura 3.5.

Usualmente, para trabalhar com predição de conjuntos de imagens, a melhor abordagem é utilizar *CNNs*, em que imagens passam por camadas convolucionais, nas quais vários filtros extraem características importantes. Por outro lado, quando os dados são séries temporais, são utilizadas *RNNs*, como *LSTM* e *GRU*, onde o modelo passa do estado anterior para o próximo estado da sequência.

Porém, quando os dados são sequências multidimensionais, é possível combinar *CNNs* com *RNNs*. Um exemplo é a rede *ConvLSTM*, que determina o estado futuro de uma determinada célula no *grid* pelos *inputs* e estados passados de seus vizinhos locais. Isso é feito usando um operador de convolução nas transições de estado para estado e de *input* para estado.

Para tratar de dados financeiros, a *ConvLSTM* supera outros modelos de série temporal, apresentando o menor erro (Bloch e Book 2021).

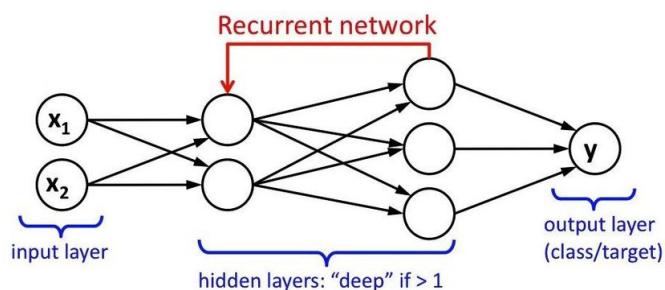
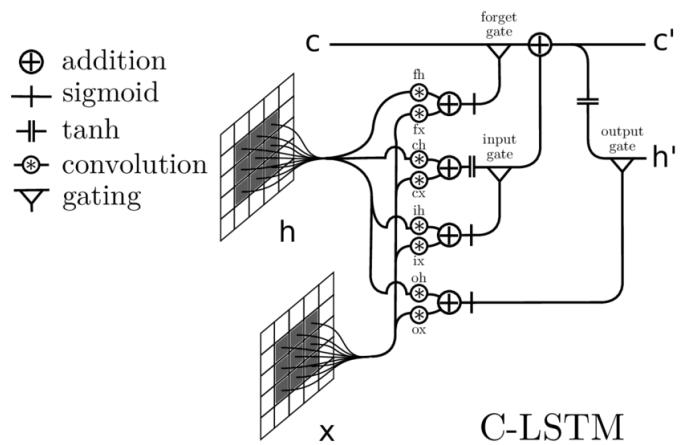


Figura 3.4: Recurrent Neural Network *Fonte: Mishra, Agarwal e Puri 2018*

Figura 3.5: ConvLSTM *Fonte: Stollenga 2016*

## Dados

### 4.1 Obtenção dos Dados

Todos os dados utilizados como base para esta monografia foram obtidos através do *website* da B3 (Bolsa do Brasil): [Preços Históricos da B3](#). Nesses arquivos estão compiladas as informações mais relevantes sobre os ativos negociados na B3, como: códigos e nomes das companhias, códigos das ações, código *ISIN*, tipo de mercado (mercado de balcão, derivativos ou mercado à vista), especificação do tipo de ação (preferencial ou ordinária), preços (preços de fechamento, abertura, mínimo, máximo, fechamento do dia anterior), volume médio negociado e outros. A B3 disponibiliza dados históricos de até 36 anos atrás.

Devido ao excesso de dados disponíveis, os arquivos de cotações históricas utilizados são disponibilizados de forma compactada. Após descompressão, os arquivos precisam ser lidos através de um layout (também disponível no *website* supracitado).

A implementação para obtenção dos dados pode ser encontrada no Apêndice 7.

	data_pregao	cod_bdi	cod_papel	tp_merc	nome_resum	espec_papel	prazo_dias_termo	moeda	preco_abertura	preco_maximo	...	preco_melhor_venda	num_negocios	qtd_titulos	vol_titulos	preco_exerc	indicador_correcao	data_vencimento
0	2015-01-02 00:00:00	2	AAPL34	10	APPLE	DRN	0	R\$	29,65	29,65	...	0,00	2,0	32300,0	958435,0	0,00	0	0000-12-31 00:00:00
1	2015-01-02 00:00:00	2	ABCB4	10	ABC BRASIL	PN EJS N2	0	R\$	12,85	12,85	...	12,56	531,0	117100,0	1469337,0	0,00	0	0000-12-31 00:00:00
2	2015-01-02 00:00:00	2	ABEV3	10	AMBEV S/A	ON EJ	0	R\$	16,14	16,58	...	16,01	21543,0	19110800,0	307341972,0	0,00	0	0000-12-31 00:00:00
3	2015-01-02 00:00:00	78	ABEVA16	70	ABEV E/J	ON	0	R\$	0,82	0,82	...	0,00	1,0	100,0	82,0	15,42	0	2015-01-19 00:00:00
4	2015-01-02 00:00:00	78	ABEVA17	70	ABEV E/J	ON	0	R\$	0,11	0,12	...	0,07	4,0	2900,0	307,0	16,92	0	2015-01-19 00:00:00
5	2015-01-02 00:00:00	78	ABEVA45	70	ABEV E/J	ON	0	R\$	1,35	1,35	...	0,00	1,0	1000,0	1350,0	15,17	0	2015-01-19 00:00:00
6	2015-01-02 00:00:00	78	ABEVA46	70	ABEV E/J	ON	0	R\$	0,37	0,37	...	0,00	1,0	1500,0	655,0	15,92	0	2015-01-19 00:00:00
7	2015-01-02 00:00:00	78	ABEVA56	70	ABEV FM/EJ	ON	0	R\$	0,49	0,53	...	0,40	46,0	177700,0	75280,0	16,17	0	2015-01-19 00:00:00
8	2015-01-02 00:00:00	78	ABEVA65	70	ABEV FM/EJ	ON	0	R\$	0,78	0,99	...	0,00	33,0	90100,0	77275,0	16,87	0	2015-01-19 00:00:00
9	2015-01-02 00:00:00	78	ABEVAT	70	ABEV E/J	ON	0	R\$	0,17	0,17	...	0,40	3,0	7500,0	1125,0	16,42	0	2015-01-19 00:00:00

Figura 4.1: Arquivo de Dados Históricos B3 (2017).

Todos os códigos utilizados estão disponibilizados no Apêndice 7.

## 4.2 Tratamento dos Dados

Após obter dados dos últimos 7 anos, foi feito um pré-processamento a fim de transformá-los em informações úteis e suficientes para obtenção das volatilidades implícitas.

O primeiro passo é filtrar os dados para os ativos de interesse. Assim, foram selecionados os preços de opções europeias para PETR4 entre 2015 e 2022.

Após filtrados, os códigos das opções foram traduzidos, através da Tabela 4.1, a fim de obter a data de vencimento de cada um dos contratos.

Mês	Call	Put
Janeiro	A	M
Fevereiro	B	N
Março	C	O
Abril	D	P
Maio	E	Q
Junho	F	R
Julho	G	S
Agosto	H	T
Setembro	I	U
Outubro	J	V
Novembro	K	W
Dezembro	L	X

Tabela 4.1: Convenção de Vencimentos *Fonte: B3 Expiry Agenda*

Um exemplo é *ABEVA16*, uma opção *call* de *ABEV3* (AMBEV S.A.), com vencimento no mês Janeiro mais próximo à data do pregão em questão.

## 4.3 Obtenção da SELIC

Além de obter os vencimentos e *strikes*, é necessário, para aplicar o *BSMI*, também obter a Taxa Livre de Risco. No caso do Brasil, a taxa em questão é a *SELIC*, a taxa básica de juros do país.

O comportamento histórico da *SELIC* pode ser encontrado na Figura 4.2.

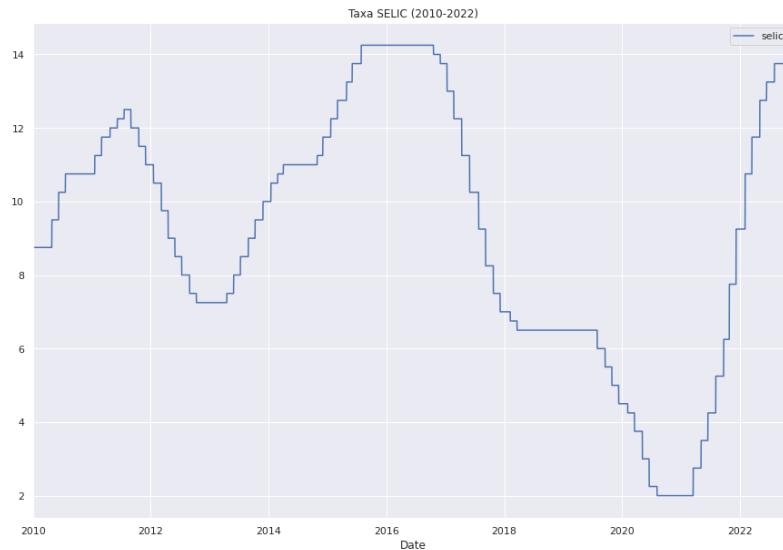


Figura 4.2: SELIC - Taxa de Juros Básica - (2010-2022). *Fonte: Própria.*

No Apêndice 7, há a implementação para obtenção da taxa histórica desde 2010.

## 4.4 Moneyness

Como mostrado em 3.2, os Sorrisos de Volatilidade são obtidos em relação aos *Moneyness* de cada uma das opções. O conceito de *Moneyness* diz respeito a uma medida relativa do *strike* em relação ao *spot* atual.

Com isso, para cada *strike*, em cada um dos tenors e a cada dia foram calculados os *Moneyness* utilizando a Equação 4.1:

$$x(K, T, t) = \ln\left(\frac{\text{Strike}}{\text{Spot} * (1 + \text{SELIC})^{\frac{(T-t)}{252}}}\right) \quad (4.1)$$

Amostras de *Log-Forward Moneyness* para opções de PETR4 podem ser encontradas na Figura 4.3.

	Ticker	Strike	SELIC	Dias até Vencimento	Spot	$x(T, K, t)$
0	PETRA1	10.66	0.1175	11	9.57	0.103016
1	PETRA10	8.61	0.1175	11	9.57	-0.110558
2	PETRA12	12.91	0.1175	11	9.57	0.294520
3	PETRA13	13.91	0.1175	11	9.57	0.369125
4	PETRA14	11.66	0.1175	11	9.57	0.192682

Figura 4.3: *Log-Forward Moneyness*. *Fonte: Própria*

## 4.5 Volatilidade Implícita

Com os *strikes*, *spots*, tempos até o vencimento, taxa *SELIC* ajustados, foi aplicado o *BSMI* e as volatilidades implícitas para cada um dos *strikes* e *tenors* foram obtidas para cada uma das datas do *dataset*.

Um exemplo de Superfície de Volatilidade obtida através do processamento dos dados da B3 pode ser encontrado na Figura 4.4.

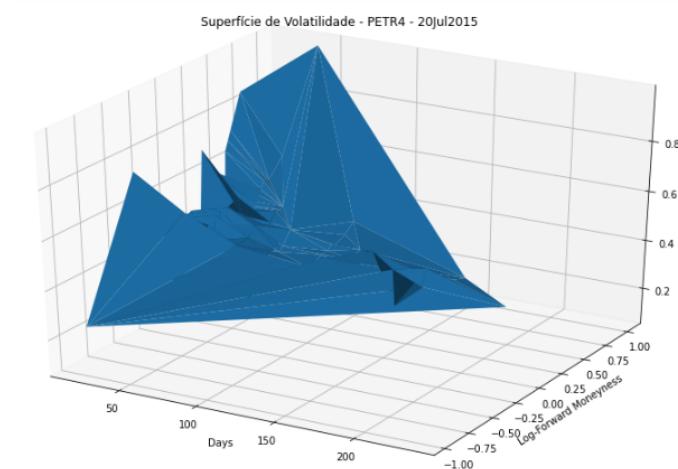


Figura 4.4: Superfície de Volatilidade - PETR4. *Fonte: Própria.*

## 4.6 Variância

Como descrito em 3.2, SVI é um modelo de parametrização para a variância total que é calculada como:

$$\Sigma^2(K, T) = (T - t) * \omega(K, T) \quad (4.2)$$

Em que  $\omega(K, T)$  é a volatilidade escalada no tempo  $T$  de cada strike  $K$ .

Na fase de processamento dos dados, após aplicar o *BSMI*, foi possível calcular a variância total para cada ponto utilizando a Equação 4.2, para então aplicar o método de Otimização Diferencial (ver 3.2.1) e obter os parâmetros de risco.

Com isso, foram obtidas as variâncias totais para cada um dos *Log-Forward Moneyness*. Na Figura 4.5, há amostras de variâncias totais calculadas utilizando 4.2.

	Ticker	$x(T, K, t)$	Vol.	Implícita	Variância	Total
0	PETRA1	0.103016		0.654427		0.018695
1	PETRA10	-0.110558		0.726080		0.023012
2	PETRA12	0.294520		0.748137		0.024432
3	PETRA13	0.369125		0.809352		0.028594
4	PETRA14	0.192682		0.678424		0.020091

Figura 4.5: Variâncias Totais - PETR4. *Fonte: Própria.*

## Modelo

### 5.1 Requisitos

Como mencionado na seção 1, o principal objetivo deste trabalho é prever a Superfície de Volatilidade Implícita (*IVS*) de ativos brasileiros do mercado de ações, seguindo a metodologia proposta por Bloch e Book 2021. Assim, as seguintes etapas propostas na seção 2 são:

- Obtenção das Volatilidades Implícitas em preços de mercado de opções através do método de Black and Scholes Inverso;
- Obtenção de Fatores de Risco paramétricos do modelo *SVI*;
- Obtenção de um Grid Fixo de Parâmetros;
- Predição dos Fatores de Risco utilizando *ConvLSTM*;
- Análise dos resultados através de *MAE*.

A fim de alcançar o objetivo mencionado, a metodologia mencionada deverá seguir os seguintes requisitos:

- Será capaz de prever superfícies até 30 dias a frente. Dado que este projeto é pioneiro na análise das volatilidades implícitas do mercado brasileiro, utilizar espaços de tempo maiores poderiam levar a propagação massiva de erros, assim como atrasos nos resultados. Portanto, predição para períodos maiores será objeto de estudos futuros;
- Apresentará um modelo capaz de prever a volatilidade de ativos brasileiros de alta liquidez (com alto volume de negociação) como PETR4, VALE3, ITUB4, BBDC4, BBAS3 e o Índice IBOVESPA;

- O erro encontrado (MAE) deverá ser similar aos de artigos como Medvedev (2019) e Bloch e Book (2021). Erros maiores podem significar especificidades do mercado brasileiro;
- O modelo será capaz de prever superfícies entre os *strikes* e *tenors* do *grid* fixo, mencionados na seção 5.2;
- Os resultados deverão seguir as regras de não-arbitragem definidas na seção 3.2.

## 5.2 Dinâmica da IVS

No Mercado de derivativos, para um determinado dia  $t$ , existem diversas opções com diferentes vencimentos  $T$  (conjunto de dimensão  $J_t$ ) e diferentes *strikes*  $K$  de (conjunto de dimensão  $I_t$ ). Como mencionado na seção 3.1, as volatilidades para cada conjunto  $(K, t, T)$  são encontradas através do *BSMI* e o modelo *SVI* é *fittado* para encontrar um conjunto de parâmetros de dimensão  $(a, b, \rho, \sigma_{em})$  para cada  $t$  e *tenor*. A fim de padronizar as dimensões dos conjuntos  $K$  e  $T$ , é criado um *grid* fixo de *strikes* e *tenors* de dimensões  $I \times J$ , através de interpolação linear das variâncias totais, obtendo por fim, um conjunto de parâmetros a cada tempo  $t$ , que pode ser denotado por:

$$M_t = \{M_{1,t}^\tau, \dots, M_{P,t}^\tau\} \quad (5.1)$$

Em que:

$$M_{p,t}^\tau = \{M_{p,t}^{T_1-t}, \dots, M_{p,t}^{T_J-t}\} \quad (5.2)$$

Neste trabalho, os *tenors* escolhidos são 7, 14, 21, 30, 60 dias. Portanto,  $J = 5$  e, dada a superfície paramétrica,  $P=5$ .

Assim, para cada tempo  $t$ , existe um conjunto de parâmetros de dimensões (5.5).

### 5.2.1 Interpolação

Como mencionado na seção 5.2, foi necessário criar um *grid* fixo de volatilidades, a fim de padronizar a dimensão do conjunto de parâmetros a cada tempo  $t$ . Para isso, foi aplicada uma interpolação linear nos dados. Essa interpolação linear foi feita para as volatilidades e possibilitou obter os parâmetros interpolados. Com isso, foi possível traçar o comportamento de cada um deles através do tempo. Um exemplo está descrito na imagem 5.1.

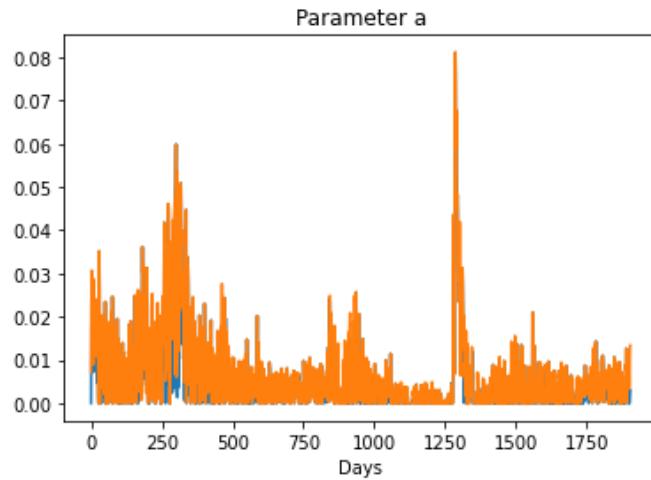


Figura 5.1: Exemplo de parâmetro (a) ao longo do tempo para  $tenor = 7$ . *Fonte: Própria.*

Além disso, tendo os parâmetros interpolados, é possível obter os Sorrisos interpolados, como na Figura 5.2.

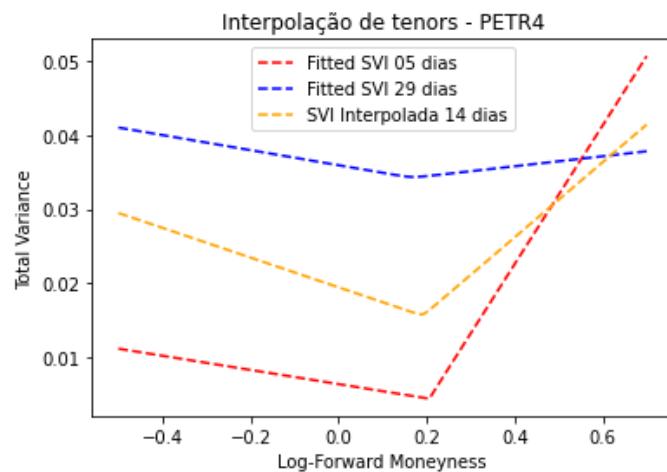


Figura 5.2: Sorriso de Volatilidade Interpolado. *Fonte: Própria.*

Através, portanto, da interpolação linear, foram gerados os Sorrisos para o *grid* fixo de *tenors*, ilustrados parcialmente na Figura 5.3.

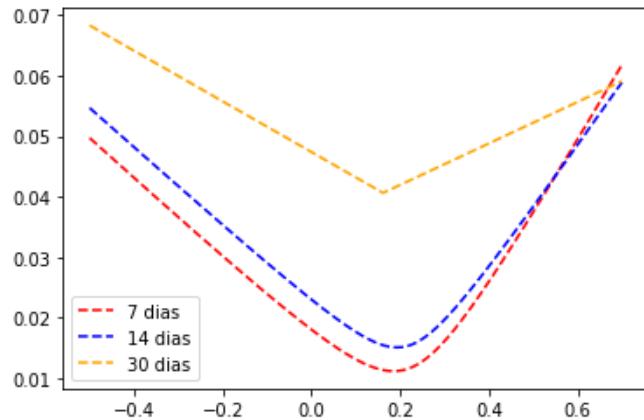


Figura 5.3: Sorriso de Volatilidade Tenors Fixos. *Fonte: Própria.*

## 5.3 Parâmetros da Rede

### 5.3.1 ConvLSTM

É essencial especificar como fazer previsões de m-passos à frente ao usar *RNN*. Decidiu-se usar uma técnica que prevê todas as etapas futuras de uma só vez, igualando o número de níveis LSTM/ConvLSTM ( $l$ ) ao número de etapas ( $m$ ). Como a abordagem pode produzir qualquer número de previsões no futuro sem exigir uma modificação completa na arquitetura, pode simplesmente se utilizar a saída de cada célula LSTM/ConvLSTM do decodificador como a previsão da *IVS*.

### 5.3.2 Inputs e Outputs

Observadas no tempo  $t$ , as variáveis explicativas  $x_i$ , são um histórico de todos os parâmetros do modelo alterados para os  $n$  períodos de tempo anteriores, bem como últimas  $n$  estruturas de termos de preços futuros a termo para os  $J$  vencimentos. Isto é:

$$\xi_t = \{M_t, M_{t-1}, \dots, M_{t-n+1}, F_t^\tau, F_{t-1}^\tau, \dots, F_{t-n+1}^\tau\} \quad (5.3)$$

Como resultado, no tempo  $t$ , o número total de variáveis explicativas é  $n * (P + 1) * J$ . Simplificando,  $x = (x_1, \dots, x_n)$  representa a sequência de variáveis explicativas, com cada membro  $x_i, i = 1, \dots, n$  tendo dimensão  $d = (P + 1, J)$ . A dinâmica da *IVS* é construída através do aprendizado da dinâmica das variáveis explicativas usando um modelo codificador-decodificador. Um codificador recorrente com unidades *LSTM* convolucionais (*ConvLSTM*).

A dimensão  $d$  de cada uma das variáveis explicativas é representada pela matriz de tamanho  $(P + 1, J)$ , e o tensor que representa o sistema dinâmico será de tamanho

$(n, d)$ . Dessa forma, o modelo (*ConvLSTM*) é capaz de compreender as relações espaço-temporais entre *strikes* e vencimentos.

A série de vetores de saída  $y = (y_1, \dots, y_T)$  será calculada dada uma sequência de vetores de entrada  $x = (x_1, \dots, x_T)$  com período  $T$ , uma pilha de  $N$  camadas ocultas ligadas recorrentemente, as sequências de vetores ocultos  $h = (h_{n,1}, \dots, h_{n,T}), n = 1, \dots, N$ . Cada vetor de saída  $y_t$  é utilizado para parametrizar uma distribuição preditiva  $P(x_{t+1}|yt)$  sobre o conjunto de entradas seguinte  $x_{t+1}$ . Equações iterativas de  $t = 1$  até  $T$  e de  $n = 2$  até  $N$  são usadas para calcular as ativações de camada oculta.

Os estados iniciais da rede de predição e as saídas das células são replicados a partir do último estado da rede de codificação. Toda a sequência de *input* é comprimida em um tensor de estado oculto  $H$  pela codificação *LSTM*, e a previsão final é fornecida pela previsão *LSTM* desdobrando esse estado oculto. Todos os elementos de *input* e saída são tensores 2D, portanto, toda a informação espacial é preservada. A rede possui muitas camadas *ConvLSTM*, o que resulta em alto poder de representação ideal para previsões em sistemas dinâmicos complicados.

Neste trabalho os hiperparâmetros escolhidos para a rede são:

- tamanho do *kernel*: (1,1);
- redes com 64 neurônios;
- taxa de dropout de 2.5
- 1 camada de convolução;
- 3 camadas ocultas;

Já a arquitetura consiste em:

- *Encoder-Decoder* com 2 filtros *ConvLSTM2D* de dimensão 8 e normalização em lote;
- Um encoder simples com 2 camadas *ConvLSTM2D*, com 64 neurônios por camada, normalização em lote;
- camada final *Conv2D* .

## 5.4 Modelo Benchmark - Naive Model

Assim como Bloch e Book (2021), uma possível implementação para quantificar a qualidade das previsões obtidas é comparar o modelo proposto com um modelo de *benchmark*. Neste trabalho, foi utilizado um modelo conhecido como "modelo ingênuo" (*Naive Model*),

cuja metodologia consiste em utilizar os dados atuais como as previsões para o próximo período. É um modelo que demanda pouco esforço e pouca manipulação, além de ser considerado um dos mais clássicos e simples para predição de valores temporais. O modelo utilizado consiste em um *random walk*, cuja definição, segundo Santos (1997), é: *”a melhor previsão para amanhã são os acontecimentos de hoje”*.

## 5.5 Cálculo de Erro

Para cada previsão feita é calculado o erro de teste da *IVS* comparado à volatilidade realizada do mercado para o dia equivalente, além de calcular a diferença entre o modelo proposto e o modelo *naive*. As performances dos modelos de previsão e de *benchmark* são mensuradas para todo o grupo de validação e a sua análise é feita através do *MAE*, que é calculado como:

$$\sum_{i=1}^D |x_i - y_i| \quad (5.4)$$

# CAPÍTULO 6

---

## Resultados

Nesta seção serão apresentados os resultados obtidos nas duas etapas chaves da metodologia:

- Otimização dos Sorrisos de Volatilidade;
- Predição das Superfícies ao longo do tempo.

### 6.1 Comportamento dos Fatores de Risco

Foram obtidos os fatores de risco através da metodologia da seção 3.2 e da criação do *grid* fixo mencionado 5.2.1 para cada dia do conjunto de dados. Os parâmetros para 04/01/2010 podem ser encontrados na Figura 6.1.

		a	b	rho	sigma	m
	date dias					
2010-01-04 00:00:00	7	0.0029	0.0021	-0.3547	0.0246	0.3287
	14	0.0043	0.0062	-0.3388	0.0359	0.2907
	21	0.0066	0.0132	-0.3110	0.0557	0.2243
	30	0.0098	0.0253	-0.3761	0.0677	0.1507
	60	0.0170	0.0575	-0.7000	0.0706	0.0084

Figura 6.1: Exemplo de parâmetros para 04/01/2010 *Fonte: Própria.*

Na Figura 6.2 estão os parâmetros ajustados obtidos e seus comportamentos ao longo do tempo.

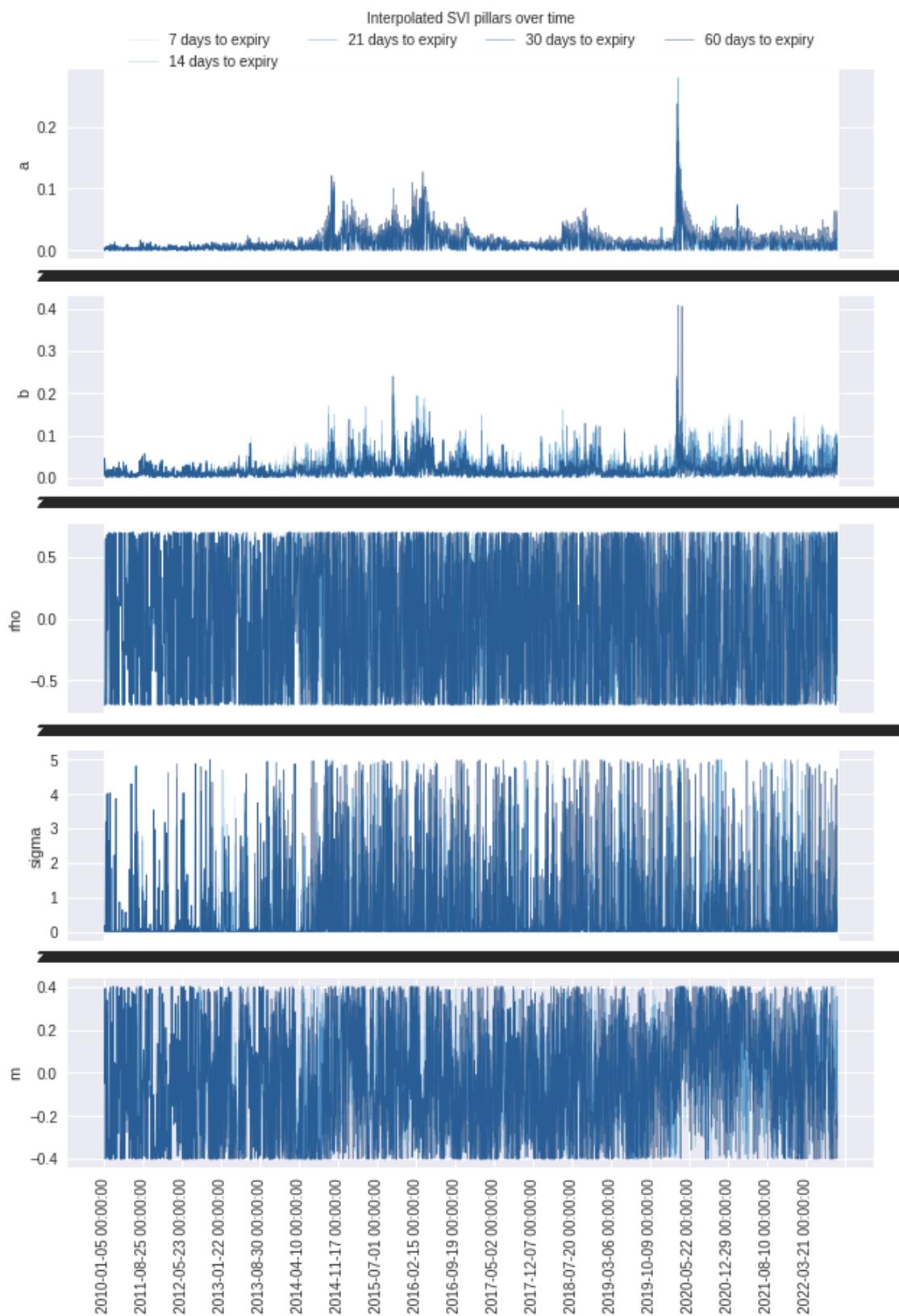


Figura 6.2: Comportamento dos fatores de risco *Fonte: Própria.*

Os resultados demonstram comportamentos esperados dos parâmetros  $a$ ,  $b$  e  $m$ . Por outro lado, os parâmetros  $\sigma$  e  $\rho$  foram responsáveis por demonstrar dificuldades no ajuste do modelo para dias com poucos *strikes* ou poucos *tenors*. O resultado dos dois últimos parâmetros mencionados indica possíveis limitações que o *SVI* pode vir a encontrar no ajuste de Sorrisos do mercado brasileiro.

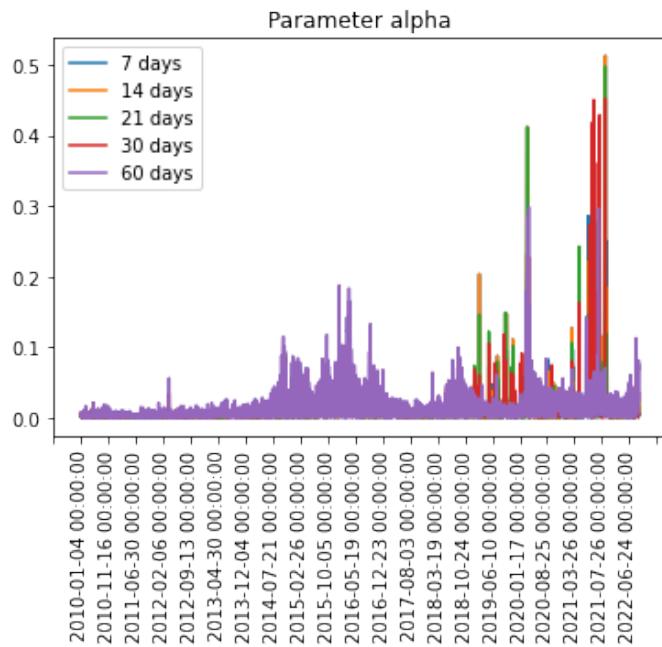


Figura 6.3: Comportamento de  $\alpha$  ao longo do tempo *Fonte: Própria.*

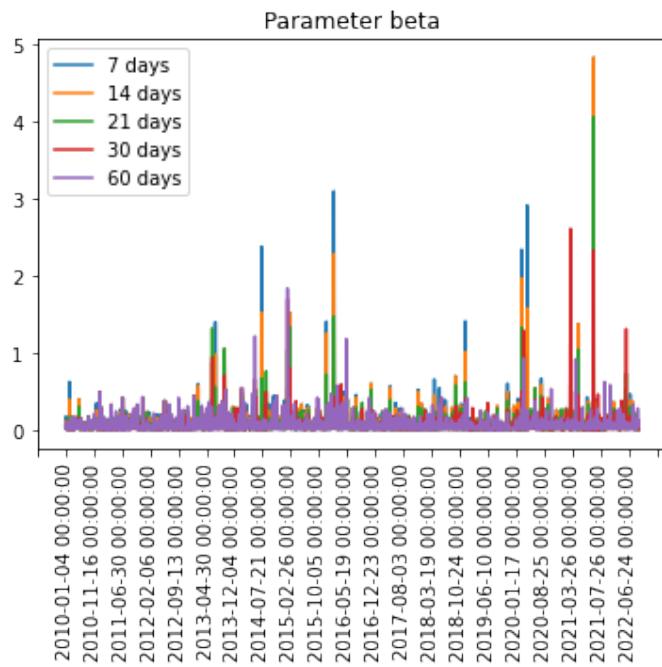


Figura 6.4: Comportamento de  $\beta$  ao longo do tempo *Fonte: Própria.*

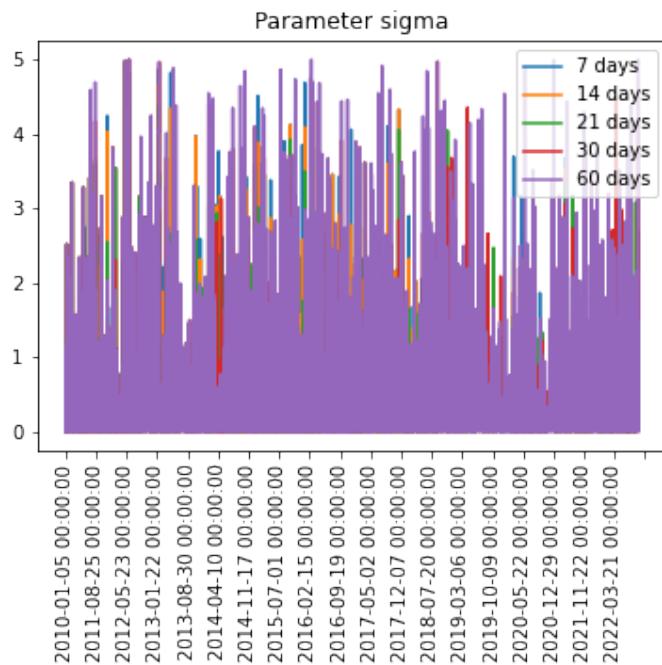


Figura 6.5: Comportamento de  $\sigma$  ao longo do tempo *Fonte: Própria.*

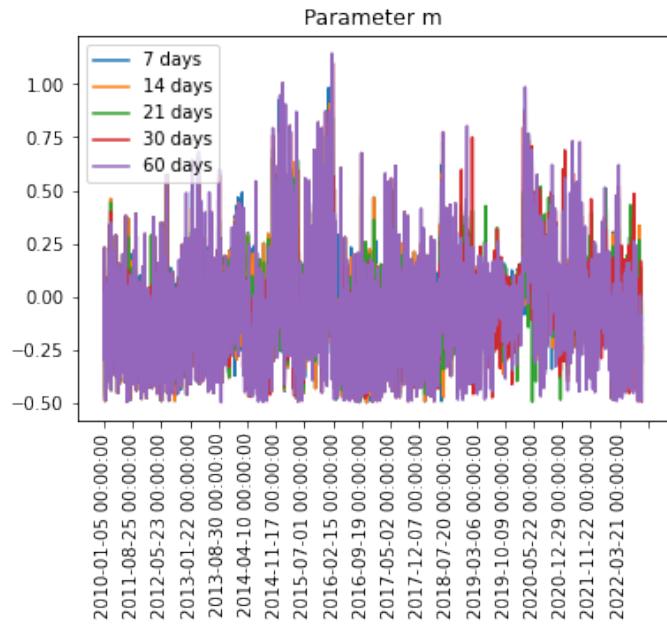


Figura 6.6: Comportamento de  $m$  ao longo do tempo *Fonte: Própria.*

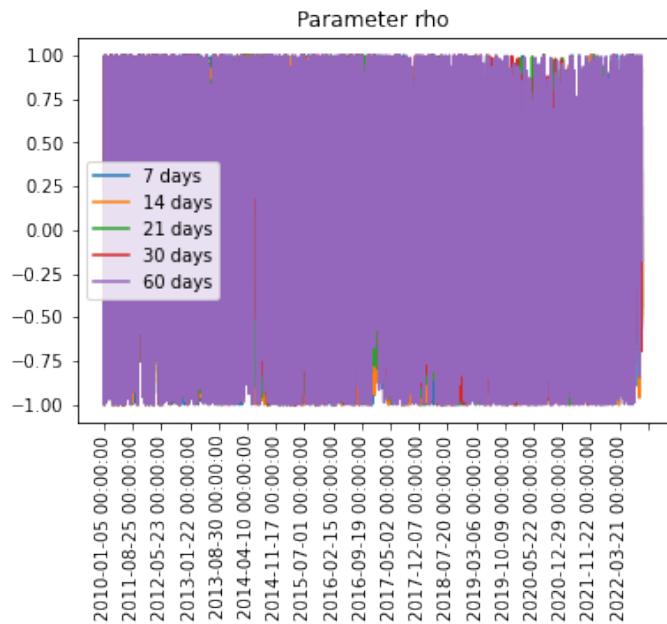


Figura 6.7: Comportamento de  $\rho$  ao longo do tempo *Fonte: Própria.*

## 6.2 Predição dos Fatores de Risco

Utilizando os períodos de treino (05-01-2010, 31-05-2016) e de teste (01-06-2016, 09-09-2022), foi treinada a rede, cujos parâmetros estão descritos na 6.8. Os erros médios para a

predição podem ser encontrados na Tabela 6.1 e os erros médios relativos ao *benchmark* estão na Tabela 6.2. As discussões sobre os resultados encontrados se encontram na seção 6.3.

```
Model: "model"

Layer (type)          Output Shape         Param #
=====
input_1 (InputLayer)    [(None, None, 5, 5, 1)]  0
conv_lstm2d (ConvLSTM2D) (None, 5, 5, 64)      16896
batch_normalization (BatchN (None, 5, 5, 64)      256
ormalization)
conv2d (Conv2D)         (None, 5, 5, 1)        65
reshape (Reshape)       (None, 5, 5)           0
=====
Total params: 17,217
Trainable params: 17,089
Non-trainable params: 128

None
Training data goes from: 2010-01-26 00:00:00 to 2016-12-29 00:00:00
  Train input shape:  (1420, 10, 5, 5, 1)
  Train target shape: (1420, 5, 5)
Test data goes from: 2017-01-02 00:00:00 to 2022-09-09 00:00:00
  Test input shape:   (1409, 10, 5, 5, 1)
  Test target shape: (1409, 5, 5)
```

Figura 6.8: Parâmetros de treinamento. *Fonte: Própria.*

Na Figura 6.9, está a convergência para as *MAEs* nos dados de teste e treino.

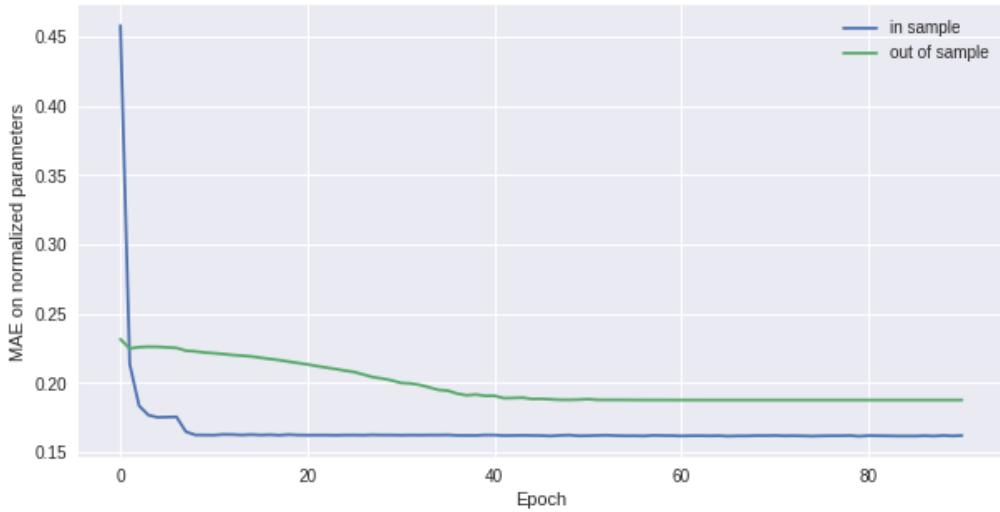


Figura 6.9: Convergência das MAEs. Fonte: Própria.

Os erros médios obtidos foram:

Tenor	t+1 MAE	t+3 MAE	t+5 MAE	t+10 MAE	t+20 MAE	t+30 MAE
7	0.1782	0.2019	0.2174	0.2605	0.3871	0.5103
14	0.1188	0.1295	0.1413	0.1846	0.2940	0.3937
21	0.0929	0.1033	0.1146	0.1571	0.2550	0.3432
30	0.0728	0.0870	0.1009	0.1442	0.2335	0.3159
60	0.0516	0.0641	0.0760	0.1115	0.1784	0.2340

Tabela 6.1: MAE no dataset de treino (05-01-2010, 31-05-2016) e de teste (01-06-2016, 09-09-2022)

Já os erros médios relativos ao modelo *naive* foram:

Tenor	t+1 MAE	t+3 MAE	t+5 MAE	t+10 MAE	t+20 MAE	t+30 MAE
7	0.0120	0.0115	0.0054	0.0057	0.1401	0.2212
14	-0.0093	-0.0087	-0.0066	0.0164	0.1079	0.1962
21	-0.0090	-0.0071	-0.0007	0.0195	0.0955	0.1752
30	-0.0024	-0.0001	0.0075	0.0309	0.1032	0.1773
60	0.0085	0.0097	0.0152	0.0355	0.0886	0.1362

Tabela 6.2: MAE Relativa (Model - Naive) no dataset de treino (05-01-2010, 31-05-2016) e de teste (01-06-2016, 09-09-2022)

Exemplos de predições podem ser encontrados nas Figuras 6.10 e 6.11. Na primeira, há a predição ao longo do tempo do valor *at-the-money* da volatilidade. Já na segunda, há um exemplo para o dia 07/05/2018, prevendo 5 dias a frente comparado com o Sorriso realizado no dia 11/05/2018 e com a predição pelo modelo *naive*.

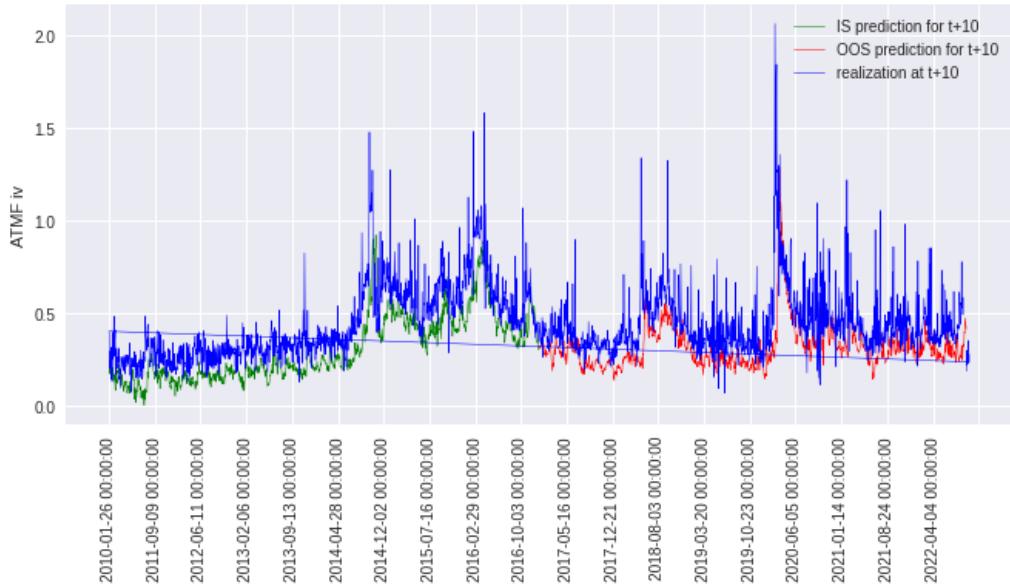


Figura 6.10: Predição de valores *at-the-money*. *Fonte: Própria.*

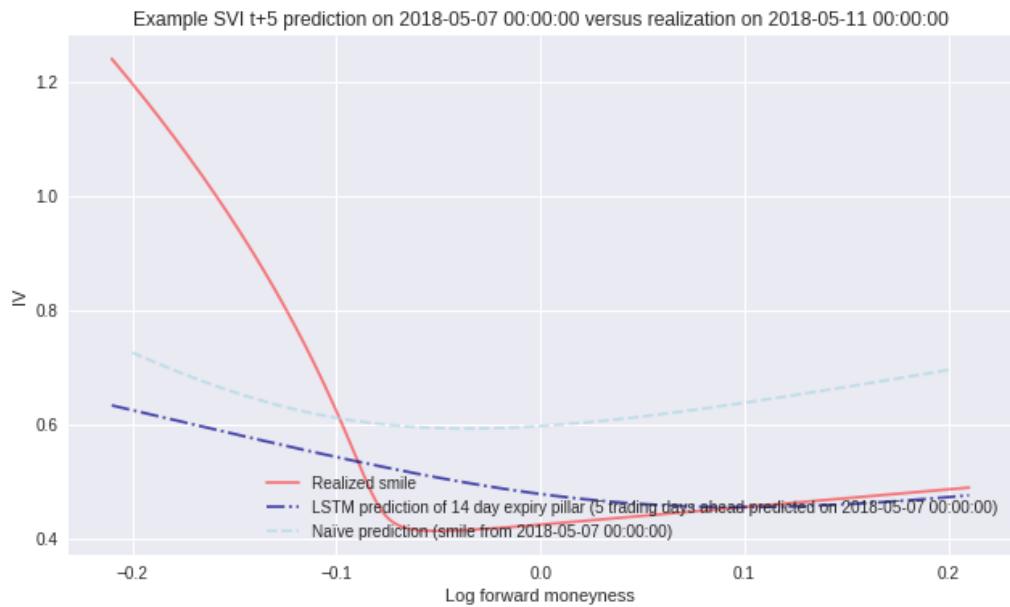


Figura 6.11: Exemplo de Sorriso previsto vs Realizado e *Naive* para o dia 07/05/2018, prevendo 5 dias à frente. *Fonte: Própria.*

## 6.3 Análise dos Resultados

Dados os resultados obtidos na seção 6.2, é possível analisar que a performance do método supera o *benchmark* para predição em curtos espaços de tempo (entre t+1 e t+10) e para opções com vencimentos de médio termo (entre 14 e 30).

Diferentemente dos resultados apresentados por Bloch e Book (2021), que concluíram que para o Índice S&P as previsões utilizando a metodologia mencionada performavam acima do modelo *naive* para vencimentos de longo prazo e para espaços de previsão maiores que  $t+3$ . Essa divergência nas conclusões utilizando modelos semelhantes pode ser justificada pelos seguintes fatores:

- Inerentes à realidade brasileira
  1. Elevada taxa de juros no Brasil;
  2. Ausência de liquidez em *tenors* mais longos;
  3. Instabilidades políticas e econômicas (Risco-Brasil).
- Ativos subjacentes diferentes
  1. Fatores Micro (*Singles stocks*).

### 6.3.1 Inerentes à realidade brasileira

#### Elevada taxa de juros

Um dos possíveis fatores que explicariam os resultados obtidos com o treinamento é a elevada taxa de juros no Brasil. A taxa de juros é utilizada tanto na precificação por *BSM*, quanto na criação do *grid* fixo com os Log-Forward Moneyness.

Para ilustrar o efeito que a *SELIC* tem nos resultados das previsões o ativo PETR4, por exemplo, em 31 de outubro de 2014, possuía um preço igual a 6,78 reais e a *SELIC* era de 11,25%, considerando um *strike* a 4,15 com vencimento em 60 dias, o cálculo do  $\eta_{Brasil}$  resulta em -0,52, logo seria descartado por estar nas condições de não-arbitragem, porém ao se considerar as mesmas condições para a taxa de juros americana (0,9%) o cálculo de  $\eta_{USA}$  resulta em -0,49, ou seja, não seria descartado.

Além disso, o modelo *BSM* considera como hipótese que a taxa de juros é constante, o que não acontece na realidade brasileira (ver Figura 6.12). Portanto, além dos resultados apresentarem as volatilidades implícitas na curva, há um erro considerável que surge pela variação da taxa "livre de risco" brasileira. Para solucionar essa distorção gerada pela variação da taxa *SELIC*, um mecanismo comum é utilizar o nível do futuro da taxa negociado na B3 para cada vencimento considerado. Essa implementação será escopo de trabalhos futuros. Já no cenário americano, usado em Bloch e Book 2021, o comportamento da taxa de juros é mais próxima de um cenário livre de risco, estressando apenas em conjunturas de crises globais.



Figura 6.12: SELIC vs Fed Funds. *Fonte: Própria.*

### Ausência de Liquidez em vencimentos longos

Diferentemente da oferta de opções de *S&P*, a liquidez das opções sobre ações brasileiras varia com o *tenor*: *tenors* mais longos possuem volume menor de negociação, bem como menor número de *strikes* disponíveis. Um exemplo pode ser encontrado na Figura 6.13, em que são apresentados os Sorrisos de Volatilidade de PETR4 para 20/04/2015. Na imagem é possível perceber a redução do número de *strikes* a medida que se avança nos *tenors* (pontos em vermelho na figura). Essa ausência de liquidez é justificada por: falta de apetite por riscos em opções mais longas, que possuem mais convexidades, além da instabilidade política e econômica que são características do mercado brasileiro.

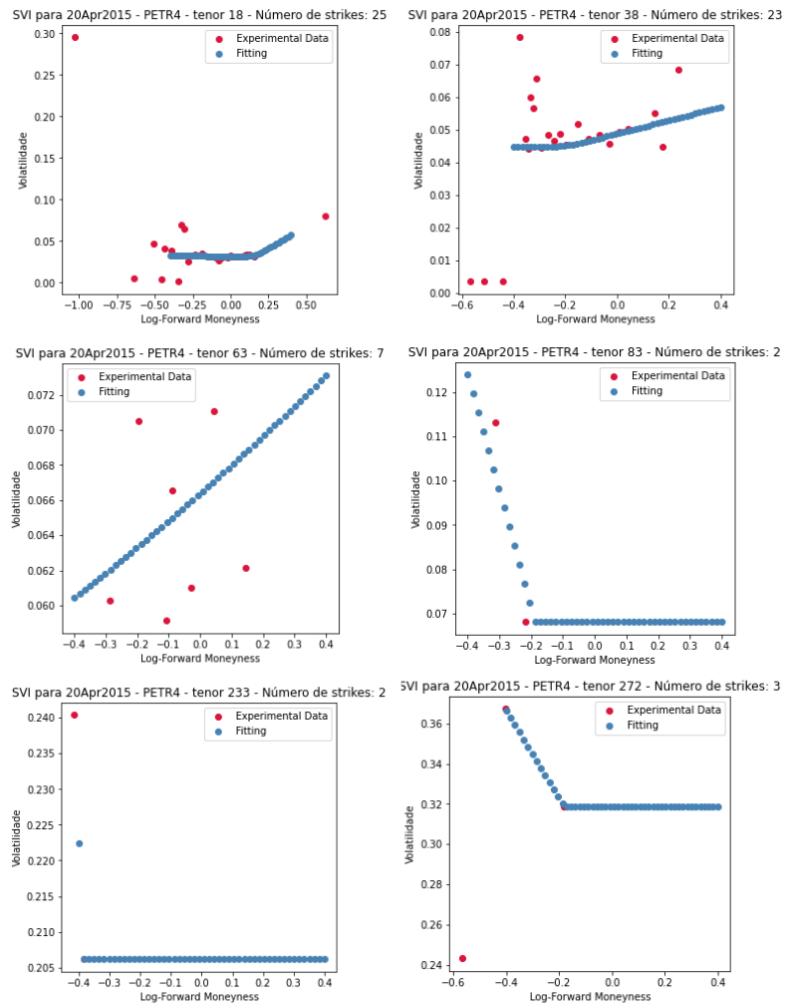


Figura 6.13: Número de *strikes* por *tenor*. Fonte: Própria.

### Risco-Brasil

Além dos fatores supracitados, é necessário considerar as diferenças implícitas na realidade brasileira devido a ser um país emergente, de alta instabilidade econômica e política.

Essas diferenças da realidade brasileira contra a estudada por Bloch e Book (2021), nos Estados Unidos, se traduzem em níveis basais de volatilidade superiores neste trabalho, o que pode ser visto na Figura 6.14, além da taxa de juros elevada e de picos nos parâmetros de risco.

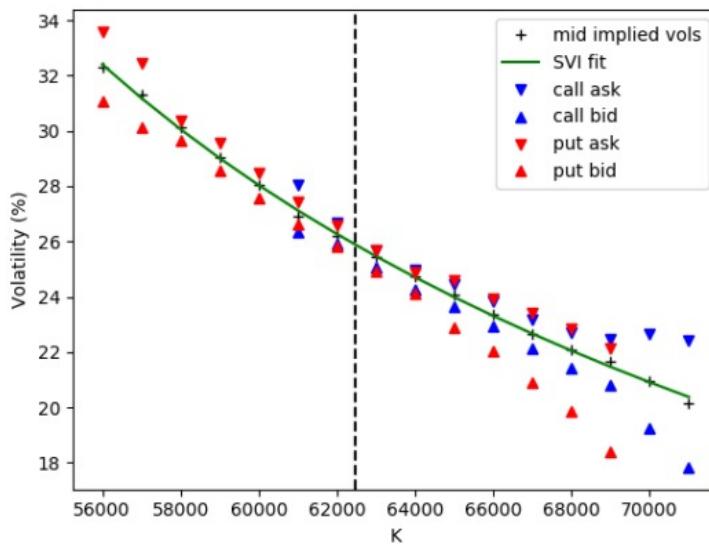


Figura 6.14: Sorriso de Volatilidade de BOVESPA. *Fonte: Yamamoto (2017).*

### 6.3.2 Ativos subjacentes diferentes

#### Fatores Micro (*Singles stocks*)

Quando prevendo superfícies de ações, fatores inerentes a acontecimentos específicos podem afetar a volatilidade. Esses fatores podem ser: publicação de resultados, anúncio de dividendos, decisões na governança corporativa que desagradam os acionistas, entre outros.

Na Figura 6.15, é possível capturar o retorno da ação PETR4 após decisões acerca do CEO serem divulgadas.

Com isso, ao comparar os erros médios obtidos neste trabalho com os apresentados por Bloch e Book (2021), cujo ativo objeto era o Índice S&P, erros maiores na predição de longos prazos são esperados, devido ao fato de opções com vencimentos mais distantes serem mais sensíveis à mudanças na Volatilidade Implícita, que, em opções sobre ações, tende a ser maior que em Índices.

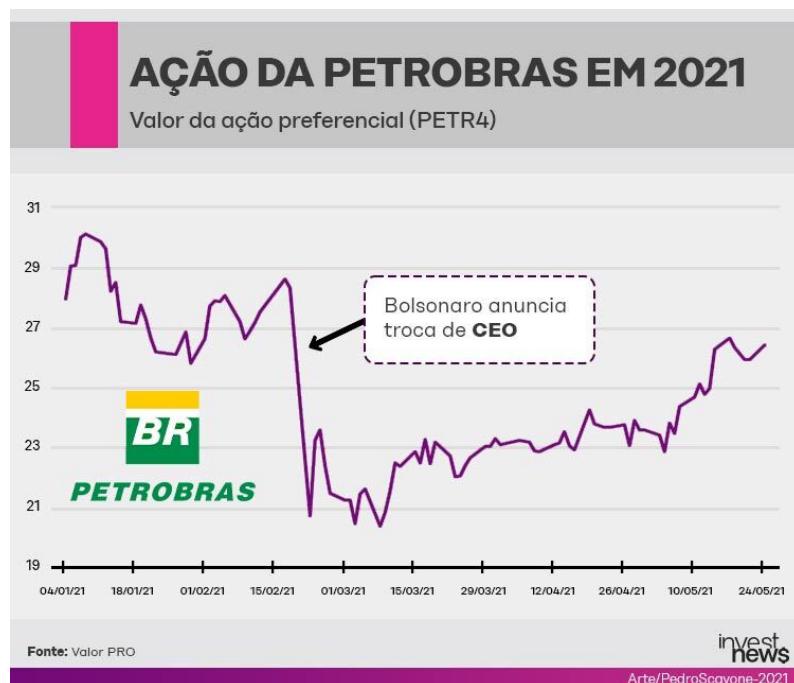


Figura 6.15: Efeito de notícias. *Fonte: Valor e InvestNews.*

---

## Conclusão

Como mencionado, este trabalho é pioneiro na aplicação de técnicas de *Deep Learning* na predição de *IVS* de ativos do mercado brasileiro de ações. Nele foi utilizado o modelo paramétrico e estocástico denominado *SVI*, que já havia se provado útil para o cenário brasileiro por Yamamoto (2017). Com os fatores de risco obtidos através do *fitting* de valores reais de variância total ao modelo *SVI*, fatores de risco foram obtidos para cada tempo  $t$ . Esses fatores foram então utilizados para treinamento de uma rede *ConvLSTM* codificada, capaz de prever, para curto prazo, a *IVS* de PETR4.

Apesar do estudo ter sido realizado utilizando opções de Petrobrás (PETR4), é esperado que o método traga resultados próximos para opções de outros ativos considerados de alta liquidez da B3, sendo eles: BBAS3, VALE3, ITUB4, BBDC4 e o próprio Índice IBOVESPA.

O modelo se mostrou eficaz na predição dos Sorrisos com vencimento em curto prazo (*tenors* inferiores a 30 dias) e para curtos períodos de tempo (até  $t+5$  dias). Estudos posteriores devem ser feitos a fim de aprimorar os resultados para longos períodos. As principais limitações encontradas no modelo estão relacionadas à ausência de liquidez em opções de vencimentos longos, à elevada taxa de juros brasileira (que tem impacto direto na especificação das opções), à instabilidade política e econômica do país, além de eventos internos de extremo impacto no retorno dos ativos.

Para continuidade dos estudos acerca da predição de *IVS* brasileiras, podem ser propostos como próximos passos:

- Testes com variações do modelo *SVI*, como por exemplo o *SSVI* e o *SVI natural*;
- Utilização de outros modelos paramétricos, como o *SABR*;
- Alterações nos parâmetros da rede, adicionando mais camadas convolucionais;
- Testes com diferentes ativos brasileiros;

- Remoção dos dados históricos eventos pontuais em que as volatilidades ficaram estressadas, como por exemplo o ano de 2020, em que a pandemia do *Covid-19* gerou distorções nos retornos de ativos globalmente.

---

## Referências

- Ahoniemi, K. (2006). "Modeling and Forecasting Implied Volatility-an Econometric Analysis of the VIX Index". Em: *Helsingin yliopiston digitaalinen arkisto*.
- Audrino, F. e D. Colangelo (2011). "Forecasting Implied Volatility Surfaces". Em: *SSRN Electronic Journal*.
- Aurell, A. (2014). "The SVI implied volatility model and its calibration". Em: *KTH SCI*.
- Bernales, A. e M. Guidolin (2014). "Can we forecast the implied volatility surface dynamics of equity options? Predictability and economic value tests". Em: *Journal of Banking and Finance*.
- Black, F. e M.S. Scholes (1973). "The pricing of options and corporate liabilities". Em: *Journal of Political Economy Volume 81 Number 3*.
- Bloch, D. e Arthur Book (2021). "Deep Learning Based Dynamic Implied Volatility Surface". Em: *SSRN Electronic Journal*.
- Brent, Richard P (2013). *Algorithms for minimization without derivatives*. Courier Corporation.
- Chen, S. e Z. Zhang (2020). "Forecasting Implied Volatility Smile Surface via Deep Learning and Attention Mechanism". Em: *SSRN Electronic Journal*.
- Chen, S., Z. Zhou e S. Li (2016). "An efficient estimate and forecast of the implied volatility surface: A nonlinear Kalman filter approach". Em: *Economic Modelling*.
- Cheng, S. L. e C. Hwang (2001). "Optimal approximation of linear systems by a differential evolution algorithm". Em: *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, v. 31, n. 6, pp. 698-707*.
- Cont, R. e F. Paris (fev. de 2002). "Dynamics of Implied Volatility Surfaces." Em: *Quantitative Finance 2 45*.
- Dumas, B., J. Fleming e R.E. Whaley (1996). "Implied volatility functions: empirical tests". Em: *The Journal of Finance, 53: 2059-2106*.
- Dupire, B. (1994). "Pricing with a smile". Em: *Risk, 7(1), 18-20*.

- Gatheral, Jim e Antoine Jacquier (2014). "Arbitrage-free SVI volatility surfaces". Em: *Quantitative Finance* 14.1, pp. 59–71. DOI: [10.1080/14697688.2013.819986](https://doi.org/10.1080/14697688.2013.819986).
- Hull, J. (2009). *Options, Futures and Other Derivatives*. Prentice Hall finance series. Pearson/Prentice Hall. ISBN: 9780136015864. URL: <https://books.google.com.br/books?id=sEmQZoHoJCcC>.
- Hull, J., W. Suo e T. Daglish (2001). "Volatility Surfaces: Theory, Rules of Thumb and Empirical Evidence". Em: *Quantitative Finance*, 7:5, 507-524,
- Huynh, S. (2018). "Modelling and Forecasting Implied Volatility Using Neural Network". Em: *Hanken School of Economics*.
- Jackwerth, J.C. e M. Rubinstein (1996). "Recovering probabilities from option prices". Em: *The Journal of Finance*, 51: 1611-1631.
- Medvedev, N. (2019). "Multi-Step Forecast of the Implied Volatility Surface Using Deep Learning". Em: *Journal of Futures Markets*, 42, 645– 667.
- Mishra, V., S. Agarwal e N. Puri (jan. de 2018). "Comprehensive and Comparative Analysis of Neural Network". Em: *INTERNATIONAL JOURNAL OF COMPUTER APPLICATION* 2.
- Rubinstein, M. (1994). "Implied binomial trees". Em: *The Journal of Finance*, 49: 771-818.
- Santos, J.E. (1997). "Previsão de volatilidade no Brasil: RiskMetrics, GARCH, volatilidade implícita ou uma combinação desses modelos? Um estudo empírico". Em: *FGV EAESP - CDEE: Teses, Doutorado em Economia de Empresas* [46].
- Shi, X. et al. (mai. de 2015). "Convolutional LSTM network: A machine learning approach for precipitation nowcasting". Em: *Advances in Neural Information Processing Systems* 28.
- Stollenga, M. (mai. de 2016). "Advances in Humanoid Control and Perception". Em: *Thèse de doctorat: Università della Svizzera italiana*, 2016.
- Storn, R. e K. Price (1997). "Differential Evolution - a Simple and Efficient Heuristic for Global Optimization over Continuous Spaces". Em: *Journal of Global Optimization*.
- Sun, K. (2018). "Using Machine Learning Methods to Predict Implied Volatility Surfaces for SPX Options". Em: *Operations Research and Financial Engineering*, 2000-2022.
- Wang, J. et al. (2017). "Modelling the implied volatility surface based on Shanghai 50ETF options". Em: *Economic Modelling*.
- Yamamoto, Rubens Y. (2017). "Estudo do método SVI aplicado à construção da volatilidade implícita para opções de ação e de índice no mercado brasileiro". Em: *Dissertação (Mestrado em Matemática, Estatística e Computação) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo*. DOI: [doi:10.11606/D.55.2018.tde-19022018-144817](https://doi.org/10.11606/D.55.2018.tde-19022018-144817).

Zhu, Y. (2018). "Comparison of Three Volatility Forecasting Models". Em: *Doctoral dissertation, The Ohio State University*.

---

# Apêndices

## Apêndice I - Códigos

Repositório com códigos utilizados

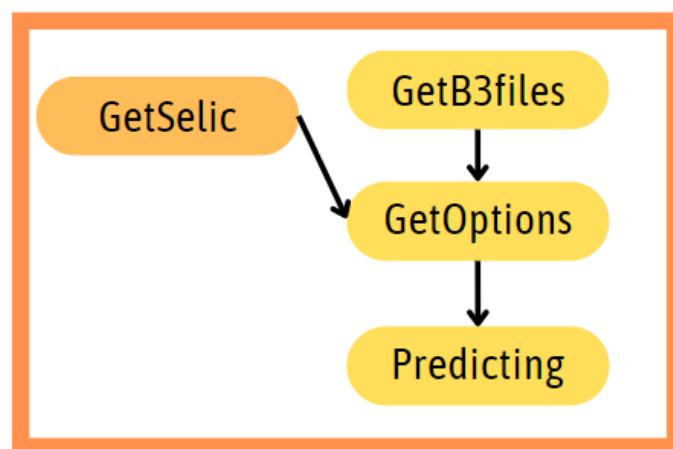


Figura A.1: Estrutura de Arquivos. *Fonte: Própria.*

---

# Anexos

## Anexo A - Put BSM

$$P(S_t, K, t) = Ke^{-r(T-t)}\Phi(-d_2) - S_t\Phi(-d_1) \quad (7.1)$$

$$d_1 = \frac{1}{\sigma\sqrt{t}} \left[ \ln\left(\frac{S}{K}\right) + t\left(r + \frac{\sigma^2}{2}\right) \right] \quad (7.2)$$

$$d_2 = \frac{1}{\sigma\sqrt{t}} \left[ \ln\left(\frac{S}{K}\right) + t\left(r - \frac{\sigma^2}{2}\right) \right] \quad (7.3)$$

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}z^2} dz \quad (7.4)$$

## Anexo B - Condições de Não-Arbitragem

Gatheral e Jacquier (2014) provaram a necessidade de limitar os parâmetros do modelo SVI de forma que ele fosse construído livre de oportunidades de arbitragem. Com isso, foram adotadas premissas no que tange a variância total do modelo:

- Deve ser linear nas *wings*, com um ponto central de curvatura;
- Deve ser positiva;
- $0 \leq a + b\sigma\sqrt{1 - \rho^2}$
- Tem relação linear com o *Log-Forward Moneyness*.

Além disso, outra condição necessária é  $|\delta\omega(K, T)| \leq 4$ , para todo *strike*  $K$  (Yamamoto 2017).

Com as condições acima satisfeitas, garante-se que a IVS é livre de arbitragens de calendário, de *butterfly* ou de qualquer outro tipo de arbitragem.