



UNIVERSIDADE FEDERAL DE OURO PRETO - UFOP
ESCOLA DE MINAS
DEPARTAMENTO DE ENGENHARIA DE MINAS - DEMIN



ISABELA PEREIRA SANTIAGO SILVA

**PREDIÇÃO DA RESISTÊNCIA DE FILITOS EM FUNÇÃO DA ANISOTROPIA POR
MEIO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL**

Ouro Preto – MG
2026

ISABELA PEREIRA SANTIAGO SILVA

isabela.santiago@aluno.ufop.edu.br

**PREDIÇÃO DA RESISTÊNCIA DE FILITOS EM FUNÇÃO DA ANISOTROPIA POR
MEIO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL**

Monografia submetida à apreciação da banca examinadora de graduação em Engenharia de Minas da Universidade Federal de Ouro Preto, como parte dos requisitos necessários para a obtenção de grau de Bacharel em Engenharia de Minas.

Orientador: Prof. Dr. Allan Erlikhman Medeiros Santos



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
ESCOLA DE MINAS
DEPARTAMENTO DE ENGENHARIA DE MINAS



FOLHA DE APROVAÇÃO

Isabela Pereira Santiago Silva

Predição da resistência de filitos em função da anisotropia por meio de métodos de inteligência artificial

Monografia apresentada ao Curso de Engenharia de Minas da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Engenharia de Minas

Aprovada em 23 de fevereiro de 2026

Membros da banca

Doutor - Allan Erlikhman Medeiros Santos - Orientador - Universidade Federal de Ouro Preto
Doutora - Tatiana Barreto dos Santos - Universidade Federal de Ouro Preto
Doutor - Hernani Mota de Lima - Universidade Federal de Ouro Preto

[Allan Erlikhman Medeiros Santos, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 24/02/2026]



Documento assinado eletronicamente por **Allan Erlikhman Medeiros Santos, PROFESSOR DE MAGISTERIO SUPERIOR**, em 24/02/2026, às 08:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1063342** e o código CRC **6ED6FE8C**.

Dedico este trabalho aos meus pais, pelo apoio incondicional, e a todos que acreditaram na minha capacidade de transformar esforço em conquista.

AGRADECIMENTOS

Primeiramente, agradeço a Deus pela força e perseverança ao longo desta trajetória acadêmica, especialmente nos momentos de maior desafio.

À minha família, em especial ao meu pai, Marcelo, e à minha mãe, Claudia, à minha irmã, à minha madrinha, ao tio Júlio e à tia Eliane, pelo apoio incondicional, incentivo constante e por sempre acreditarem em mim. Cada conquista é também fruto do amor e da confiança que recebi ao longo dessa caminhada.

Ao meu orientador, Allan Erlikhman, pela orientação técnica, disponibilidade e contribuições essenciais para o desenvolvimento deste trabalho. Agradeço pela confiança, discussões construtivas e direcionamento científico que enriqueceram esta pesquisa.

À professora Rosa Malena, pelos anos de Iniciação Científica e por acreditar na minha dedicação. Aos professores do curso de Engenharia de Minas da UFOP, pelas orientações e ensinamentos que contribuíram significativamente para minha formação acadêmica e profissional.

Ao Hemerson, pela coorientação e confiança no meu trabalho, e ao Luiz, técnico do laboratório, pelo apoio nas atividades desenvolvidas. Ao Henrique, pelo apoio e pela ajuda durante o processo de escrita deste TCC, tornando essa etapa mais leve e organizada.

À Universidade Federal de Ouro Preto, pela estrutura, oportunidades e ambiente acadêmico que possibilitaram meu crescimento pessoal e profissional.

À Minera Júnior e à Liga Acadêmica ARGUS, pelas experiências em gestão, planejamento e trabalho em equipe, que contribuíram para minha formação como engenheira.

Ao Pablo, pelo apoio constante, pelas palavras de incentivo e por acreditar na minha força ao longo desta jornada.

Aos amigos e colegas de graduação, pela parceria ao longo dessa jornada. À República Toka, pelas vivências e pelo suporte durante essa etapa tão intensa da vida acadêmica. Em especial, à Ana Luiza, por compartilhar comigo os desafios e as conquistas desse processo.

Por fim, agradeço a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

“Aprender é a única coisa de que a mente nunca se cansa, nunca tem medo e nunca se arrepende.”

Leonardo da Vinci

RESUMO

A resistência de pico de rochas anisotrópicas é fortemente influenciada pela orientação da foliação e pela tensão confinante, apresentando comportamento direcional não linear. Modelos clássicos de resistência fornecem resultados pontuais associados a orientações específicas e demandam elevado número de ensaios laboratoriais, dificultando a representação contínua do comportamento mecânico. Nesse contexto, este trabalho teve como objetivo interpolar a resistência de pico de filitos anisotrópicos por meio da aplicação e comparação de modelos preditivos baseados em técnicas de aprendizado de máquina supervisionado. Foram analisados cinco bancos de dados experimentais compilados da literatura, representando os litotipos *Carbonaceous*, *Quartzite*, *Micaceous*, *Dry* e *Moretown*. A modelagem foi implementada em linguagem Python utilizando os algoritmos *Random Forest*, *Gaussian Process Regression* e *Support Vector Regression*. O desempenho foi avaliado por meio das métricas: coeficiente de determinação, erro médio absoluto e raiz do erro quadrático médio, considerando conjuntos de treinamento e teste. Os resultados confirmaram a presença de anisotropia direcional em todos os litotipos, com curvas características em formato típico de “U” e redução relativa do contraste anisotrópico com o aumento da tensão confinante. Entre os modelos avaliados, o *Gaussian Process Regression* apresentou o melhor desempenho global, combinando elevados valores de coeficiente de determinação e menores erros relativos para a maioria dos litotipos, enquanto o *Support Vector Regression* destacou-se na representação do filito *Micaceous*. O modelo *Random Forest* demonstrou capacidade preditiva satisfatória, embora sem se sobressair em relação aos demais. Conclui-se que modelos de aprendizado de máquina supervisionado são capazes de representar de forma contínua e coerente o comportamento mecânico anisotrópico de filitos, constituindo ferramenta promissora para complementar métodos experimentais tradicionais.

Palavras-chave: Anisotropia; Resistência de pico; Filitos; Aprendizado de máquina; Modelagem preditiva.

ABSTRACT

The peak strength of anisotropic rocks is strongly influenced by foliation orientation and confining stress, exhibiting nonlinear directional behavior. Classical strength criteria provide discrete results associated with specific orientations and require a large number of laboratory tests, which limits the continuous representation of mechanical behavior. In this context, this study aimed to interpolate the peak strength of anisotropic phyllites through the application and comparison of supervised machine learning predictive models. Five experimental datasets compiled from the literature were analyzed, representing the Carbonaceous, Quartzite, Micaceous, Dry, and Moretown phyllite lithotypes. The modeling was implemented in Python using the Random Forest, Gaussian Process Regression, and Support Vector Regression algorithms. Model performance was evaluated using the coefficient of determination, mean absolute error, and root mean square error, considering both training and testing datasets. The results confirmed the presence of directional anisotropy in all lithotypes, characterized by typical “U-shaped” strength curves and a relative reduction in anisotropic contrast with increasing confining stress. Among the evaluated models, GPR presented the best overall performance, combining high R^2 values and lower relative errors for most lithotypes, while SVR showed superior performance in representing the Micaceous phyllite. The Random Forest model demonstrated satisfactory predictive capability, although without outperforming the other approaches. It is concluded that supervised machine learning models are capable of continuously and coherently representing the anisotropic mechanical behavior of phyllites, constituting a promising tool to complement traditional experimental methods.

Keywords: Anisotropy; Peak strength; Phyllites; Machine learning; Predictive modeling.

LISTA DE FIGURAS

Figura 1: Curvas experimentais de resistência em função da orientação β para filitos: (a) resistência à compressão uniaxial; (b) envelope de resistência para planos de fraqueza únicos e múltiplos.	24
Figura 2: Esquema genérico do modelo de <i>Random Forest</i>	26
Figura 3: Algoritmo SVM: maximização da margem de separação entre classes através de vetores de suporte e um hiperplano.	28
Figura 4: Fluxograma das etapas da metodologia do presente trabalho.	34
Figura 5: Gráficos de σ_1 vs β_{eff} para (a) <i>Carbonaceous Phyllite</i> , (b) <i>Quartzite Phyllite</i> , (c) <i>Micaceous Phyllite</i> , (d) <i>Phyllites dry</i> e (e) <i>Moretown Phyllite</i>	46
Figura 6: Gráfico do grau de anisotropia por nível de σ_3 , analisados para (a) <i>Carbonaceous Phyllite</i> , (b) <i>Quartzite Phyllite</i> , (c) <i>Micaceous Phyllite</i> , (d) <i>Phyllites dry</i> e (e) <i>Moretown Phyllite</i>	47
Figura 7: Gráfico de dispersão que mostra a relação entre a diferença de tensões ($\sigma_1 - \sigma_3$) e o ângulo efetivo (β_{eff}) para diferentes níveis de tensão confinante (σ_3).	48
Figura 8: Comparação entre valores reais e preditos de σ_1 pelo modelo <i>Random Forest</i>	53
Figura 9: Curvas $\sigma_1 - \sigma_3 \times \beta_{eff}$ preditas pelo <i>Random Forest</i>	54
Figura 10: Comparação entre valores reais e preditos de σ_1 pelo modelo GPR.	57
Figura 11: Curvas $\sigma_1 - \sigma_3 \times \beta_{eff}$ preditas pelo GPR.	59
Figura 12: Curva de aprendizado (<i>Learning Curve</i>) para o modelo GPR.	60
Figura 13: Comparação entre valores reais e preditos de σ_1 pelo modelo SVR.	63
Figura 14: Curvas $\sigma_1 - \sigma_3 \times \beta_{eff}$ preditas pelo SVR.	65
Figura 15: Curva de aprendizado (<i>Learning Curve</i>) para o modelo SVR.	66

LISTA DE TABELAS

Tabela 1: Resumo de etapas, objetivos e técnicas aplicadas.	33
Tabela 2: Lista dos filitos que basearam o banco de dados, incluindo referência bibliográfica.	35
Tabela 3: Nomenclatura das variáveis do banco de dados.	35
Tabela 4: Resumo das principais características dos bancos de dados utilizados.	44
Tabela 5: Índice de anisotropia (A) por litotipo e nível de σ_3	49
Tabela 6: Tamanhos dos conjuntos de treinamento e teste utilizados na avaliação dos modelos.	50
Tabela 7: Desempenho do modelo <i>Random Forest</i> na predição da resistência de pico (σ_1).	51
Tabela 8: Desempenho do modelo <i>Gaussian Process Regression</i> na predição da resistência de pico (σ_1).	55
Tabela 9: Desempenho do modelo SVR na predição da resistência de pico (σ_1).	61
Tabela 10: Síntese comparativa dos modelos.	67

SUMÁRIO

1	INTRODUÇÃO.....	12
1.1	Formulação do Problema	13
1.2	Justificativa	14
1.3	Objetivos.....	15
1.3.1	Objetivo Geral	15
1.3.2	Objetivos Específicos	15
1.4	Estrutura do Trabalho	16
2	REVISÃO BIBLIOGRÁFICA	18
2.1	Anisotropia em Rochas Metamórficas.....	18
2.2	Critérios Clássicos de Resistência para Rochas Anisotrópicas	18
2.2.1	Critério de Mohr–Coulomb modificado para anisotropia	18
2.2.2	Critério de Jaeger.....	19
2.2.3	Abordagens por critérios isotrópicos adaptados.....	20
2.3	Mecânica Experimental dos Filitos	21
2.3.1	Estrutura microestrutural e mineralógica do filito.....	21
2.3.2	Efeitos da orientação β na resistência.....	22
2.3.3	Influência da granulação e coesão	23
2.3.4	Modelos experimentais relevantes para filitos	23
2.4	Algoritmos de Aprendizado Supervisionado para Interpolação da Resistência	24
2.4.1	<i>Random Forest</i>	25
2.4.2	<i>Gaussian Process Regression</i> (GPR)	27
2.4.3	<i>Support Vector Machines</i> (SVM).....	28
2.4.4	Comparações entre os três métodos	29
2.5	Inteligência Artificial Aplicada à Resistência de Rochas.....	30
2.5.1	Histórico da aplicação de IA na Mecânica das Rochas	30
2.5.2	Variáveis normalmente empregadas.....	31
3	METODOLOGIA.....	32
3.1	Banco de dados	35
3.1.1	Origem e Natureza dos Dados	35
3.1.2	Variáveis de Entrada e Saída	36
3.1.3	Processamento e Variáveis Derivadas	36
3.1.4	Origem e Caracterização dos Dados de Filito	36
3.2	Pré-processamento dos Dados	37
3.3	Treinamento dos Modelos	38
3.3.1	<i>Random Forest</i>	38
3.3.2	<i>Gaussian Process Regression</i>	39
3.3.3	<i>Support Vector Regression</i>	39
3.4	Avaliação do Desempenho dos Modelos.....	40
3.4.1	Avaliação estatística	40
3.4.2	Estatísticas Descritivas do Banco de Dados	41

3.4.3	Métricas de Validação e Diagnóstico	41
3.4.4	Avaliação física e coerência mecânica	42
3.4.5	Comparação entre os Modelos	42
4	RESULTADOS E DISCUSSÕES.....	44
4.1	Apresentação dos bancos de dados e caracterização geral	44
4.2	Análise exploratória do comportamento anisotrópico	44
4.2.1	Comparação do comportamento anisotrópico entre os diferentes filitos	49
4.3	Desempenho dos modelos de interpolação da resistência de pico.....	50
4.3.1	Avaliação do modelo <i>Random Forest</i>	50
4.3.2	Avaliação do modelo de <i>Gaussian Process Regression</i> (GPR)	55
4.3.3	Avaliação do modelo de <i>Support Vector Regression</i> (SVR).....	61
4.4	Comparação entre os modelos preditivos	66
4.5	Síntese dos principais resultados	68
5	CONSIDERAÇÕES FINAIS	70
5.1	Sugestões para Trabalhos Futuros	71
	APÊNDICES	75

1 INTRODUÇÃO

A resistência mecânica das rochas é um parâmetro fundamental para a análise da estabilidade de estruturas como taludes, fundações de barragens e escavações subterrâneas. Em maciços rochosos anisotrópicos, como os filitos, essa resistência varia significativamente conforme a orientação das descontinuidades em relação à direção da tensão principal maior. A desconsideração dessa dependência direcional pode comprometer a segurança das obras geotécnicas, resultando em avaliações imprecisas da estabilidade estrutural.

A caracterização do comportamento mecânico de rochas transversalmente isotrópicas requer a determinação da resistência de pico (σ_1) por meio de ensaios de compressão triaxial, realizados sob diferentes ângulos de orientação (β) e tensões de confinamento (σ_3). Entretanto, a obtenção desses dados é frequentemente limitada por restrições de tempo, custo e escassez de amostras adequadas. A necessidade de múltiplos ensaios para capturar os diferentes graus de anisotropia torna o processo laboratorial extensivo e oneroso (Cerrato, 2024; Nasser, Rao e Ramamurthy, 2003; Singh *et al.*, 2015).

Diversos pesquisadores propuseram modificações e extensões de critérios clássicos de resistência para incorporar a anisotropia estrutural. Contudo, muitos desses modelos apresentam limitações práticas relevantes, exigindo elevado número de ensaios para determinação de parâmetros com difícil interpretação física ou baixa aplicabilidade em projetos geotécnicos. Além disso, tais abordagens nem sempre conseguem representar simultaneamente a variação direcional da resistência e a não linearidade observada experimentalmente. Para fins práticos, torna-se desejável que os modelos sejam simples, demandem número reduzido de ensaios e utilizem parâmetros com significado físico bem definido (Singh *et al.*, 2015).

No caso específico dos filitos, o comportamento mecânico é fortemente controlado pela foliação. Ainda assim, os estudos voltados a essa litologia permanecem relativamente escassos quando comparados a outras rochas anisotrópicas, como ardósias e folhelhos. Observa-se também predominância de investigações voltadas à estimativa pontual da resistência, com menor ênfase na construção de modelos capazes de representar de forma contínua o comportamento mecânico em função da orientação estrutural e do confinamento. Apesar dos avanços experimentais e teóricos, ainda não há modelos amplamente consolidados capazes de representar de forma contínua e com quantificação de incerteza o comportamento anisotrópico de filitos.

Diante dessa lacuna, este trabalho propõe o desenvolvimento de modelos de aprendizado supervisionado para a interpolação da resistência de pico de filitos anisotrópicos, utilizando dados experimentais compilados da literatura técnica. A pesquisa concentra-se na construção de superfícies contínuas de resistência em função da orientação β e da tensão confinante σ_3 . Busca-se, por meio de abordagens como a *Gaussian Process Regression* (GPR), não apenas a estimativa média da resistência, mas também a quantificação da incerteza associada, contribuindo para uma representação mais robusta e fisicamente coerente do comportamento de filitos em projetos de engenharia.

1.1 Formulação do Problema

A determinação da resistência de pico (σ_1) é um aspecto central da mecânica das rochas, sendo essencial para a análise da estabilidade de estruturas geotécnicas. Em rochas anisotrópicas, essa resistência não depende apenas do nível de confinamento, mas também da orientação dos planos de anisotropia em relação à direção da tensão principal maior, o que impõe dificuldades adicionais à caracterização do comportamento resistente.

Os critérios de resistência clássicos e suas adaptações para rochas anisotrópicas, embora relevantes do ponto de vista teórico, apresentam limitações significativas quando aplicados à prática de engenharia. Em geral, esses modelos exigem um elevado número de ensaios laboratoriais para a determinação de parâmetros, muitos dos quais possuem interpretação física limitada e dependem fortemente da orientação estrutural e do estado de tensões. Essas exigências restringem sua aplicabilidade, especialmente em contextos nos quais os dados experimentais disponíveis são escassos.

Como consequência, projetos geotécnicos frequentemente baseiam-se em informações pontuais de resistência, associadas a orientações específicas ensaiadas em laboratório, sem uma representação contínua do comportamento mecânico ao longo do intervalo angular relevante. Essa abordagem dificulta a avaliação adequada da variabilidade direcional da resistência e da incerteza associada às estimativas, podendo comprometer a confiabilidade das análises de estabilidade.

Diante desse cenário, surge o seguinte questionamento: como estimar de forma contínua e fisicamente coerente a resistência de pico de filitos anisotrópicos, considerando a dependência com a orientação estrutural (β) e a tensão confinante (σ_3), a partir de um conjunto

limitado de dados experimentais, incorporando a quantificação da incerteza associada às previsões?

1.2 Justificativa

A realização deste trabalho justifica-se, primeiramente, pela lacuna técnica e científica existente na modelagem da resistência de rochas anisotrópicas, em especial dos filitos. Apesar do reconhecimento da influência da orientação estrutural e da tensão confinante no comportamento da resistência dessas rochas, a maioria dos estudos disponíveis na literatura concentra-se na predição direta da resistência com condições específicas, deixando de lado a abordagem voltada para a interpolação contínua do comportamento mecânico. Além disso, a ausência de modelos que quantificam a incerteza associada às estimativas representa uma limitação relevante para a aplicação prática dos dados experimentais. Nesse contexto, o emprego de métodos de aprendizado supervisionado contribui para o avanço das análises desses dados e oferece uma avaliação mais completa da resistência anisotrópica de filitos.

Do ponto de vista social, a estabilidade de taludes, escavações subterrâneas e fundações está diretamente associada à segurança de pessoas, comunidades e infraestrutura. A utilização de estimativas pontuais de resistência, sem consideração adequada da variabilidade direcional e da incerteza envolvida, pode conduzir a decisões conservadoras excessivas ou, de forma mais crítica, a avaliações inadequadas da estabilidade estrutural. Assim, a adoção de modelos que forneçam não apenas valores médios de resistência, mas também informações sobre a confiabilidade das previsões, contribui para processos decisórios mais responsáveis e tecnicamente embasados em projetos geotécnicos.

Do ponto de vista aplicado à geomecânica, a proposta deste trabalho possui implicações diretas no dimensionamento e na análise de estabilidade de taludes e escavações subterrâneas em maciços anisotrópicos. Em projetos reais, é comum que a resistência adotada represente apenas valores obtidos para orientações específicas ensaiadas em laboratório, não contemplando de forma contínua a variabilidade direcional da rocha. A utilização de superfícies de resistência interpoladas em função da orientação estrutural (β) e da tensão confinante (σ_3) permite representar de maneira mais realista o comportamento do maciço ao longo de diferentes geometrias de escavação. Isso possibilita análises de estabilidade mais coerentes com a orientação predominante da foliação no campo, reduzindo incertezas na

definição de ângulos de talude, fatores de segurança e sistemas de suporte, contribuindo para projetos tecnicamente mais consistentes e potencialmente mais otimizados.

Sob os aspectos econômicos, a proposta deste trabalho apresenta relevância ao buscar reduzir a dependência de extensos ensaios laboratoriais, que são onerosos e demorados. A possibilidade de interpolar o comportamento da resistência a partir de um conjunto limitado de dados experimentais permite otimizar o uso das informações já disponíveis, reduzindo custos operacionais. Dessa forma, a metodologia proposta favorece soluções mais eficientes e alinhadas às práticas modernas de engenharia.

1.3 Objetivos

1.3.1 Objetivo Geral

Desenvolver e avaliar modelos preditivos baseados em aprendizado supervisionado para a interpolação contínua da resistência de pico (σ_1) de filitos anisotrópicos, em função da orientação estrutural (β) e da tensão confinante (σ_3), a partir de um banco de dados experimental limitado de ensaios de compressão triaxial compilado da literatura técnica.

1.3.2 Objetivos Específicos

Os objetivos específicos foram:

- Desenvolver modelos de aprendizado supervisionado baseados nos algoritmos *Random Forest* (RF), *Support Vector Regression* (SVR) e *Gaussian Process Regression* (GPR) para a interpolação da resistência de pico (σ_1);
- Avaliar o desempenho estatístico dos modelos desenvolvidos por meio de métricas como coeficiente de determinação (R^2), erro quadrático médio (RMSE) e erro absoluto médio (MAE);
- Verificar a coerência física das predições, analisando a capacidade dos modelos em reproduzir o comportamento anisotrópico da resistência em função da orientação β e da tensão confinante σ_3 ;
- Comparar os modelos preditivos quanto à capacidade de generalização, robustez frente à dispersão experimental e representação do comportamento direcional da resistência;

- Analisar a quantificação da incerteza associada às previsões, identificando regiões de maior e menor confiabilidade das estimativas.

1.4 Estrutura do Trabalho

O presente trabalho está estruturado em cinco capítulos principais, além das referências bibliográficas e os apêndices complementares.

Capítulo 1 – Introdução: apresenta uma breve introdução do tema, a formulação do problema, a justificativa da pesquisa, os objetivos: geral e específicos, bem como a estrutura do trabalho.

Capítulo 2 – Revisão Bibliográfica: reúne os fundamentos teóricos que embasam o estudo, abordando inicialmente a anisotropia em rochas metamórficas e os principais critérios clássicos de resistência aplicados a materiais anisotrópicos. Em seguida, são discutidos aspectos da mecânica experimental dos filitos e, por fim, os algoritmos de aprendizado supervisionado empregados neste trabalho, bem como sua aplicação na Mecânica das Rochas, utilizando-se de artigos científicos e bases de dados disponíveis na literatura atual.

Capítulo 3 – Metodologia: descreve os procedimentos adotados para a construção e análise dos modelos preditivos. São apresentados os bancos de dados utilizados, as variáveis consideradas, o pré-processamento dos dados, o treinamento dos modelos *Random Forest*, *Gaussian Process Regression* e *Support Vector Regression*, além dos critérios de avaliação estatística e de coerência física empregados na validação dos resultados, bem como a metodologia utilizada para a aquisição dos resultados.

Capítulo 4 – Estudo de Caso: apresenta e discute os resultados obtidos. Inicialmente, realiza-se a caracterização dos bancos de dados e a análise exploratória do comportamento anisotrópico dos filitos. Em seguida, avalia-se o desempenho individual de cada modelo preditivo e por fim realiza-se a comparação entre eles, culminando na síntese dos principais resultados.

Capítulo 5 – Considerações Finais: apresenta as conclusões do trabalho, destacando as principais contribuições da pesquisa e sugerindo direções para trabalhos futuros.

Referências bibliográficas: apresentam todas as referências utilizadas no trabalho.

Apêndices: reúnem materiais complementares relevantes à compreensão e reprodução da metodologia adotada.

2 REVISÃO BIBLIOGRÁFICA

2.1 Anisotropia em Rochas Metamórficas

A anisotropia mecânica consiste na variação das propriedades físicas e da resistência de uma rocha de acordo com a orientação (Donath, 1972). Esse comportamento é fortemente influenciado por planos de fraqueza ou descontinuidades, como juntas, falhas e estratificações. O fenômeno é recorrente em rochas metamórficas devido a características estruturais e geológicas, como a xistosidade e foliações.

Os filitos são anisotrópicos por apresentarem orientações preferenciais, desenvolvidas sob altas pressões e temperaturas durante o processo de metamorfismo. Como são constituídos predominantemente por minerais lamelares, cujas xistosidades e foliações criam planos de fraqueza bem definidos, suas características físicas e de resistência são diretamente influenciadas por essas estruturas.

A direção do plano de anisotropia é representada pelo ângulo β em relação ao plano da tensão principal maior (σ_1). A variação deste ângulo influencia diretamente a resistência à compressão uniaxial e as propriedades físicas da rocha (Donath, 1972).

As curvas experimentais de resistência à compressão uniaxial em rocha anisotrópica exibem um padrão e formato em “U”. Estudos demonstram que a resistência atinge valores máximos nas orientações de 0° e 90° , enquanto os valores mínimos ocorrem para ângulos intermediários (Ramamurthy, Rao e Singh, 1993a). Esse comportamento resulta na geometria côncava para cima observada nos gráficos de resistência versus o ângulo β .

Diante desse comportamento anisotrópico, foram propostos critérios de resistência capazes de representar a influência da orientação estrutural.

2.2 Critérios Clássicos de Resistência para Rochas Anisotrópicas

2.2.1 Critério de Mohr–Coulomb modificado para anisotropia

O critério de *Mohr–Coulomb* prevê a orientação do plano de ruptura em rochas isotrópicas, estabelecendo que a resistência ao cisalhamento é constituída pela coesão e pelo atrito interno do material. Em sua formulação original, o modelo assume que as propriedades mecânicas são constantes e independem da direção da tensão.

Contudo, essa teoria apresenta limitações ao lidar com rochas anisotrópicas, pois falha em considerar parâmetros de coesão que variam em função do ângulo β . Em rochas que possuem planos de fraqueza com resistências distintas, a abordagem isotrópica torna-se insuficiente por não captar as alterações decorrentes da orientação. Apesar disso, o modelo permanece como uma ferramenta essencial devido à sua robustez e simplicidade, podendo ser adaptado para representar o comportamento direcional de rochas anisotrópicas.

O critério de *Mohr–Coulomb* modificado introduz conceitos fundamentais para suprir as lacunas do modelo original. Destaca-se, nesse contexto, a variação da resistência em função do ângulo β , visto que o coeficiente de atrito demonstra baixa sensibilidade às variações estruturais. Em contrapartida, a coesão apresenta elevada sensibilidade, sendo controlada diretamente pela microestrutura e pela foliação da rocha (Donath, 1972). Essa distinção é crucial para a correta estimativa da carga de ruptura em maciços anisotrópicos. Além disso, o modelo adota a premissa teórica de que a resistência ao cisalhamento varia continuamente com a direção das tensões, o que implica a existência de um único plano preferencial de ruptura (Jaeger, 1960).

2.2.2 Critério de Jaeger

O critério de Jaeger clássico fundamenta-se na existência de dois conjuntos distintos de propriedades mecânicas: as da matriz rochosa e as do plano de fraqueza. Cada elemento possui valores próprios de resistência ao cisalhamento e coeficiente de atrito interno diferentes, de modo que a ruptura pode ocorrer preferencialmente ao longo do plano de fraqueza ou através do material intacto, a depender da orientação e do estado de tensões (Jaeger, 1960). Esse comportamento decorre da menor resistência ao cisalhamento nos planos de fraqueza em relação ao maciço preservado.

Essa abordagem é indicada para rochas com anisotropia estrutural marcante, caracterizada por foliação contínua ou planos de fraqueza dominantes, como filitos, ardósias e folhelhos. Nessas formações rochosas, a presença de superfícies preferenciais, comumente resultantes de falhas pré-existent, clivagem ou acamamento, controla o comportamento mecânico, fazendo com que a ruptura se desenvolva prioritariamente ao longo dessas estruturas geológicas.

Em contraste, o critério de *Mohr–Coulomb* modificado pressupõe que a anisotropia se distribui de forma contínua por todo o material. Nesse modelo, os parâmetros de resistência

variam gradualmente em função do ângulo β , tornando o comportamento do maciço integralmente dependente da orientação do carregamento. Essa distinção permite explicar a variabilidade da resistência com o ângulo sem a necessidade de isolar um único plano de fraqueza discreto.

2.2.3 Abordagens por critérios isotrópicos adaptados

Os critérios isotrópicos, tanto teóricos quanto empíricos, permanecem em amplo uso na prática da engenharia devido à sua consolidação e viabilidade operacional. A preferência pela adaptação desses modelos pauta-se na facilidade de aplicação e na menor demanda por dados experimentais complexos. Visto que ensaios laboratoriais que incluem diversos graus de anisotropia sob diferentes orientações de carga são custosos e demorados, a adaptação de critérios já estabelecidos apresenta-se como uma alternativa eficaz para estimar a resistência de maciços rochosos.

Nesse cenário, embora o critério de *Hoek–Brown* tenha sido concebido para meios isotrópicos, sua aplicação em contextos anisotrópicos é viável mediante ajustes metodológicos. (Gomes *et al.*, 2021) exemplifica essa abordagem ao incorporar a anisotropia de forma indireta, ajustando os parâmetros de resistência do material (σ_{ci} e m_i) ou aplicando fatores de escala ao estado de tensões para representar o comportamento médio do maciço. Diferentemente de modelos que isolam as descontinuidades, essa proposta trata a anisotropia como uma característica global, o que elimina a necessidade de assumir um plano de fraqueza ou estabelecer uma dependência analítica direta com o ângulo β .

A utilização desses modelos clássicos adaptados oferece vantagens significativas para o projeto prático, pois preserva as definições originais do critério e facilita a aplicabilidade. Ao manter a compatibilidade com softwares comerciais de geotecnia, essa lógica permite que o engenheiro utilize parâmetros determinados por ensaios globais. Assim, a robustez do modelo original é aproveitada para viabilizar análises em projetos reais de forma ágil.

Contudo, Cerrato (2024) aponta que o ajuste de tais variáveis exige cautela na interpretação, sob o risco de se negligenciar detalhes críticos da estrutura geológica. Do ponto de vista conceitual, os parâmetros ajustados para a anisotropia podem carecer de um significado físico evidente, visto que muitas vezes não resultam de medições laboratoriais diretas. Na prática, o modelo funciona como um ajuste estatístico aos dados existentes; ele se adapta aos resultados observados, mas não necessariamente prevê o comportamento mecânico em

cenários fora da amostra calibrada. Como não há um controle estrutural explícito sobre a ruptura, a calibração empírica torna-se estritamente dependente do conjunto de dados utilizado, limitando sua capacidade de generalização.

Essa dependência de calibrações empíricas e a necessidade de capturar incertezas e não linearidades, sem a imposição de parâmetros físicos artificiais, abrem caminho para novas abordagens. Nesse contexto, os métodos baseados em dados surgem como uma alternativa promissora. Ao contrário dos modelos empíricos rígidos, as técnicas de Inteligência Artificial conseguem processar variações contínuas e complexas inerentes às rochas anisotrópicas, permitindo uma representação mais fiel do comportamento do maciço sem a necessidade de simplificações excessivas.

2.3 Mecânica Experimental dos Filitos

Embora os critérios teóricos citados anteriormente forneçam a base conceitual para o estudo da anisotropia, a validação de sua eficácia depende da análise de dados experimentais. O comportamento mecânico anisotrópico é observado e mensurado primordialmente em ambiente laboratorial; nesse contexto, o filito configura-se como um dos casos representativos mais bem documentados na literatura, dada a sua foliação planar evidente.

Para esta fundamentação, adotam-se os estudos clássicos de Donath (1972) e Singh et al. (2015), cujos ensaios permitem correlacionar a orientação das estruturas geológicas à resistência da rocha sob diferentes estados de tensão. Nas subseções seguintes, detalham-se a caracterização mineralógica desses materiais e os respectivos resultados observados em laboratório.

2.3.1 Estrutura microestrutural e mineralógica do filito

Os filitos são rochas metamórficas de granulação fina, caracterizadas por uma foliação bem desenvolvida que se manifesta através do alinhamento preferencial de minerais micáceos. De acordo com as observações de Donath (1972), essa organização textural cria planos de clivagem paralelos e persistentes, que definem a estrutura lamelar da rocha. É essa disposição microscópica dos grãos que estabelece a base física para a anisotropia, uma vez que a continuidade dos planos de foliação é o que permite a identificação visual das direções preferenciais de fraqueza no material.

A composição mineralógica dessas rochas é caracterizada por uma matriz rica em minerais plácóides e lamelares, predominantemente muscovita, clorita e sericita. Esses minerais coexistem com uma fase mais rígida composta por grãos de quartzo, que se distribuem de forma variável. A proporção entre os minerais micáceos (mais compressíveis) e o quartzo (fase granular rígida) é o fator determinante para a definição da heterogeneidade mineralógica da amostra, estabelecendo o contraste entre as camadas de foliação e os grãos constituintes.

Análises em lâminas delgadas, como as discutidas por Ramamurthy, Rao e Singh, (1993), confirmam que a predominância relativa de minerais lamelares controla a simetria estrutural da rocha. As observações petrográficas revelam que o arranjo dos grãos de quartzo e o grau de continuidade das camadas micáceas definem a anisotropia física do material. Essa organização interna, observada em escala microscópica, fundamenta a descrição do filito como um meio transversalmente isotrópico, sem a necessidade de recorrer a parâmetros de resistência para sua classificação inicial.

2.3.2 Efeitos da orientação β na resistência

As investigações em filitos evidenciam que a resistência mecânica apresenta um comportamento nitidamente direcional, variando conforme a orientação da foliação em relação à tensão principal máxima (σ_1). Os valores obtidos não seguem uma tendência linear simples. Esse padrão foi demonstrado por Donath (1972) mediante ensaios triaxiais, nos quais se plotou a variação da resistência na ruptura em função do ângulo de anisotropia (β).

Os dados indicam a ocorrência de valores mínimos de resistência bem definidos quando a foliação se encontra orientada entre 30° e 45° em relação à direção da compressão. Em contrapartida, a resistência máxima manifesta-se nas orientações de 0° ou 90°. Tal comportamento é documentado de forma consistente em diferentes tipos de rochas foliadas, como ardósias e xistos, confirmando que a faixa angular de fraqueza é uma característica intrínseca desses materiais.

Além da orientação, a resposta mecânica é fortemente influenciada pelo nível de tensão confinante aplicado. Ensaios sob diferentes pressões demonstram que o incremento da tensão de confinamento tende a reduzir a anisotropia aparente da resistência. Esse processo diminui a disparidade entre as diferentes orientações, promovendo uma resposta mais homogênea do maciço, embora a natureza anisotrópica permaneça presente mesmo sob altas pressões.

2.3.3 Influência da granulação e coesão

As investigações laboratoriais indicam que a coesão da rocha é fortemente influenciada pela granulação e pela textura mineralógica. Donath (1972) demonstrou, por meio de estudos comparativos, que a coesão é o parâmetro mais sensível às variações microestruturais, especialmente em função da presença de minerais placóides. Em contrapartida, o coeficiente de atrito interno não apresenta variações significativas com a alteração da granularidade.

O impacto da textura sobre a resistência torna-se evidente ao comparar materiais de diferentes granulações. Em rochas finas e micáceas, como os filitos, a resistência é predominantemente controlada pela coesão da matriz, o que resulta em uma anisotropia mais pronunciada. Por outro lado, em rochas de granulação mais grossa, a contribuição do atrito interno torna-se mais relevante, tendência está confirmada nos materiais analisados.

Entre os parâmetros de resistência ao cisalhamento, a coesão mostra-se particularmente dependente da orientação da foliação. Conforme observado nas investigações de Donath (1972), a variação da resistência em função do ângulo β é governada primordialmente pela coesão, enquanto o coeficiente de atrito interno apresenta baixa sensibilidade direcional. Essa distinção física é fundamental para compreender a mecânica de ruptura em meios foliados, independentemente do modelo matemático adotado para sua representação.

2.3.4 Modelos experimentais relevantes para filitos

A consolidação dos dados discutidos nas seções anteriores em modelos gráficos permite sintetizar e visualizar o comportamento anisotrópico das rochas foliadas. Diferente de modelos teóricos explicativos, essas representações possuem caráter predominantemente descritivo, oferecendo um panorama claro de como a resistência varia em função da estrutura interna do material.

Nesse contexto, os trabalhos de Ramamurthy, Venkatappa Rao e Singh (1993b) constituem uma referência fundamental para a compreensão mecânica dos filitos. Por meio de uma série abrangente de ensaios de compressão uniaxial e triaxial, o autor investigou sistematicamente a influência da orientação dos planos de foliação (β) em relação à tensão principal maior (σ_1). Os resultados demonstram um padrão consistente e reproduzível, confirmando que a magnitude da variação da resistência é dependente da orientação estrutural da amostra.

Graficamente, essa variação é expressa por curvas de resistência versus orientação β , que tipicamente assumem um formato em "U". Conforme ilustrado na Figura 1a do trabalho de (Ramamurthy, Venkatappa Rao e Singh J., 1988; Ramamurthy, Venkatappa Rao, G. e Singh, 1993), os valores máximos de resistência são observados nas orientações de 0° e 90° , enquanto os mínimos ocorrem em ângulos intermediários, geralmente entre 30° e 45° . O envelope de resistência apresentado na Figura 1b sintetiza essa tendência para cenários de planos de fraqueza únicos e múltiplos, fornecendo uma representação visual direta da perda de resistência induzida pela anisotropia estrutural.

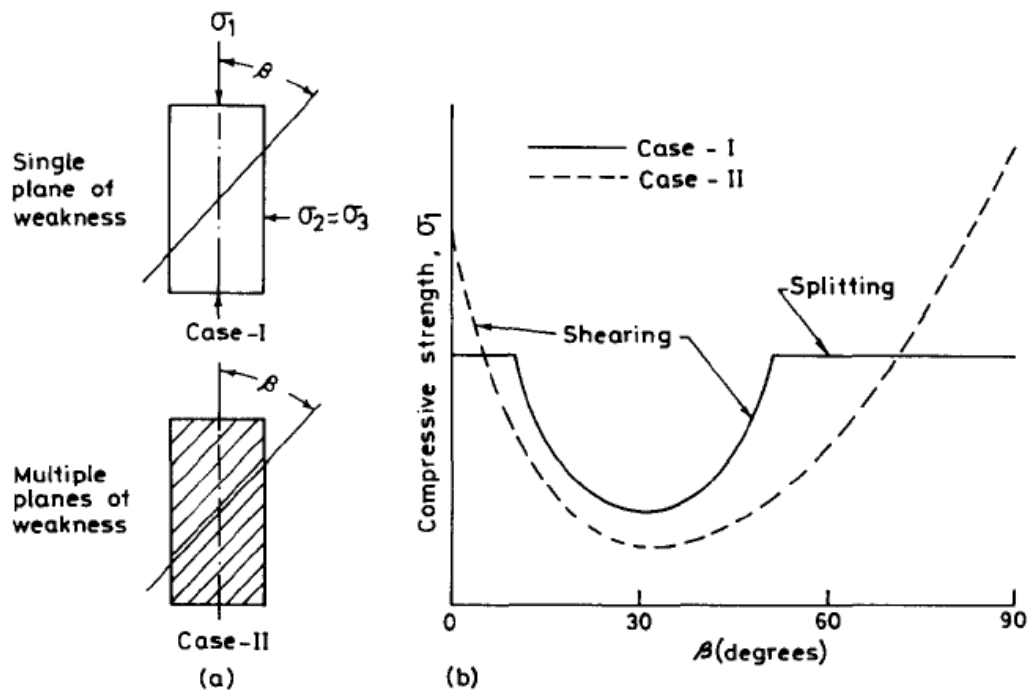


Figura 1: Curvas experimentais de resistência em função da orientação β para filitos: (a) resistência à compressão uniaxial; (b) envelope de resistência para planos de fraqueza únicos e múltiplos.

Fonte: Adaptado de (Ramamurthy, Venkatappa Rao, G e Singh, 1993).

2.4 Algoritmos de Aprendizado Supervisionado para Interpolação da Resistência

Os algoritmos de aprendizado supervisionado são aplicados em situações nas quais se dispõe de um conjunto de dados composto por variáveis denominadas preditoras, utilizadas para estimar uma variável dependente, denominada variável alvo. Nesse contexto, o algoritmo aprende a relação entre as variáveis de entrada e a variável de saída a partir de exemplos previamente observados, possibilitando a realização de previsões para novos casos (Laredo Sicsú e Lerner Barth, 2023).

Embora os modelos experimentais descrevam o comportamento anisotrópico das rochas, sua aplicação prática é frequentemente limitada pela escassez de dados e pela acentuada não linearidade das relações envolvidas. Nesse cenário, os métodos de aprendizado supervisionado surgem como ferramentas robustas para a interpolação de dados e estimativa de valores intermediários. Diferente das abordagens clássicas, esses métodos aprendem as relações diretamente das observações disponíveis, sem a imposição de uma forma funcional ou equação física pré-definida (Hastie, Tibshirani e Friedman, 2009).

Ressalta-se que o emprego desses algoritmos não substitui a compreensão física do fenômeno. Pelo contrário, o método de aprendizagem supervisionado atua de forma complementar aos resultados laboratoriais, permitindo uma representação mais contínua e detalhada da resistência anisotrópica em contextos de alta complexidade (Cerrato, 2024).

2.4.1 *Random Forest*

O *Random Forest* é um método de aprendizado supervisionado baseado na combinação de múltiplas árvores de decisão, construídas a partir de diferentes subconjuntos do conjunto de dados original. Cada árvore individual fornece uma estimativa da variável alvo, e a predição final do modelo é obtida pela média das previsões das árvores que compõem a floresta (Breiman, 2001).

A construção do modelo envolve a geração de múltiplas amostras *bootstrap* a partir do conjunto de dados original. Para cada amostra, uma árvore de decisão é crescida de forma independente, sendo que, em cada nó, apenas um subconjunto aleatório das variáveis de entrada é considerado para a escolha da melhor divisão. As árvores são crescidas sem poda, o que aumenta a variância individual, posteriormente reduzida pelo processo de agregação das previsões (Liaw e Wiener, 2002). A Figura 2 apresenta um esquema genérico do funcionamento do modelo *Random Forest*.

A introdução de diferentes fontes de aleatoriedade durante o treinamento amplia a diversidade entre os modelos individuais, reduzindo a variância global e a sensibilidade a flutuações estatísticas específicas do conjunto de dados. Como consequência, o uso de ensembles mitiga o risco de sobreajuste (*overfitting*) e confere maior estabilidade às previsões, especialmente em bases de dados reduzidas e ruidosas (Breiman, 2001; Hastie, Tibshirani e Friedman, 2009).

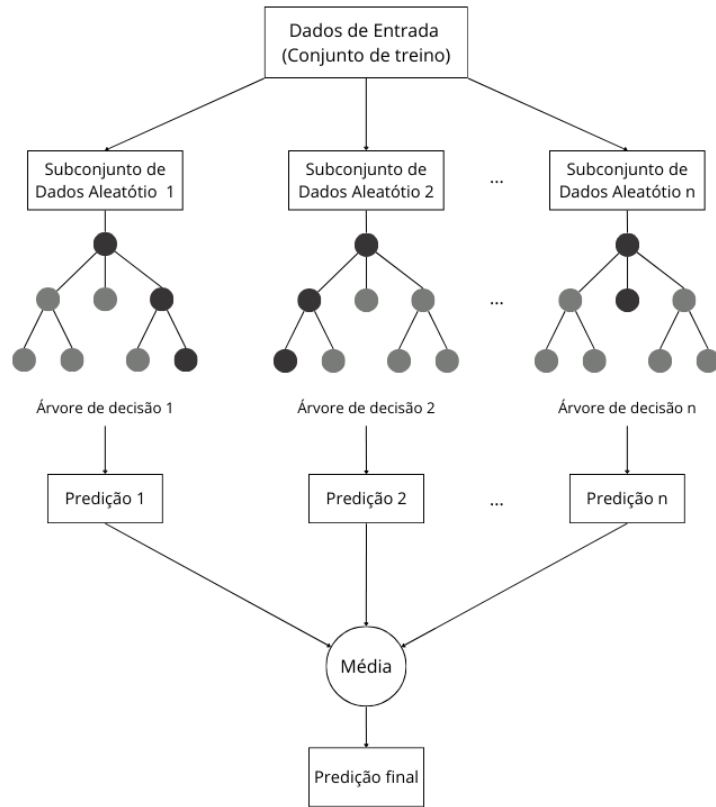


Figura 2: Esquema genérico do modelo de *Random Forest*.

Fonte: (Autora, 2026).

Essa característica torna o *Random Forest* particularmente adequado para a modelagem da resistência mecânica de rochas anisotrópicas, cujos resultados experimentais apresentam elevada dispersão. A agregação das árvores suaviza ruídos experimentais e permite capturar relações não lineares complexas, como a dependência da resistência em relação à orientação estrutural β , sem a necessidade de impor uma forma funcional prévia ou realizar ajustes exaustivos de parâmetros.

Na geotecnia, técnicas de aprendizado de máquina têm sido aplicadas para estimar propriedades mecânicas a partir de amostragens limitadas (Cerrato, 2024). Nesse cenário, o *Random Forest* é adotado como uma ferramenta estratégica de interpolação, cujo objetivo é viabilizar a construção de uma superfície de resposta contínua. Essa técnica propõe preencher as lacunas entre as orientações ensaiadas, buscando uma representação consistente com os dados disponíveis sobre o comportamento resistente do maciço rochoso.

2.4.2 *Gaussian Process Regression (GPR)*

A *Gaussian Process Regression (GPR)* é um método de aprendizado supervisionado de natureza probabilística, fundamentado em uma abordagem Bayesiana não paramétrica para inferência (Park et al., 2017). Diferentemente de modelos determinísticos, a GPR realiza previsões incorporando conhecimento prévio por meio de funções de covariância (*kernels*) e fornece, de forma explícita, medidas de incerteza associadas às estimativas realizadas (Wang, 2024).

Um processo gaussiano é definido como um conjunto de variáveis aleatórias tal que qualquer subconjunto finito dessas variáveis apresenta uma distribuição normal multivariada (Cabral e Catarina, 2021). No contexto da regressão, isso implica modelar diretamente uma distribuição conjunta sobre os valores da função desconhecida em diferentes pontos do espaço de entrada. A ideia central consiste em assumir que pontos de entrada semelhantes tendem a produzir valores de saída semelhantes, sendo essa noção de similaridade formalizada matematicamente por meio da função de covariância, ou *kernel* (Wang, 2024).

Entre as principais vantagens da Regressão por Processos Gaussianos destacam-se: (i) sua natureza não paramétrica, que evita a imposição de uma forma funcional pré-definida; (ii) a excelente capacidade de quantificação da incerteza associada às previsões; (iii) a flexibilidade para modelar relações altamente não lineares; (iv) a sólida fundamentação teórica probabilística; e (v) o bom desempenho em conjuntos de dados reduzidos (Wang, 2024).

Na formulação da GPR, a cada ponto do espaço de entrada associa-se uma variável aleatória, e a distribuição conjunta dessas variáveis é caracterizada por uma função média e por uma função de covariância (*kernel*) (Park et al., 2017). Como resultado, a predição não é expressa apenas por um valor médio esperado, mas também por uma variância preditiva, que representa explicitamente a incerteza associada à estimativa em cada ponto do domínio.

No contexto deste estudo, marcado por elevada dispersão experimental e forte variabilidade do comportamento anisotrópico das rochas, a quantificação da incerteza é tão relevante quanto a estimativa do valor médio da resistência. (Cerrato, 2024) ressalta a importância do tratamento explícito da incerteza por meio de simulações de Monte Carlo e intervalos de confiança. Diferentemente dessas abordagens, a GPR incorpora a incerteza diretamente em sua formulação probabilística, fornecendo de maneira integrada previsões pontuais e intervalos de confiança associados.

A função de covariância desempenha papel central na GPR, pois controla o grau de correlação entre os pontos do conjunto de dados e define propriedades fundamentais da função estimada, como suavidade e escala de variação (Wang, 2024). Quando se espera um comportamento contínuo e globalmente suave, como a variação da resistência mecânica em função da orientação estrutural β , o *kernel* do tipo *Radial Basis Function* (RBF) tende a produzir interpolações compatíveis com a forma geral das curvas experimentais observadas.

Em situações nas quais estão presentes irregularidades associadas à dispersão experimental, *kernels da classe Matérn* permitem um controle mais refinado do grau de suavidade da função ajustada, sem comprometer a estrutura global do fenômeno (Scikit-learn, 2026). Dessa forma, curvas de resistência em função de β , com formato típico em “U”, podem ser representadas de maneira fidedigna, reforçando a adequação da GPR para a interpolação da resistência anisotrópica em problemas envolvendo bases de dados reduzidas e elevada variabilidade experimental.

2.4.3 *Support Vector Machines* (SVM)

A *Support Vector Machine* (SVM) é um método de aprendizado supervisionado amplamente utilizado em problemas de classificação e regressão. O princípio fundamental do algoritmo consiste na determinação de um hiperplano ótimo que separa os dados no espaço de características, maximizando a margem entre as observações mais próximas desse hiperplano, denominadas vetores de suporte (Vikramaditya Jakkula, 2006). A margem corresponde à maior distância possível, paralela ao hiperplano, que não contém pontos de dados em seu interior. A Figura 3 ilustra o conceito de maximização da margem e a atuação dos vetores de suporte.

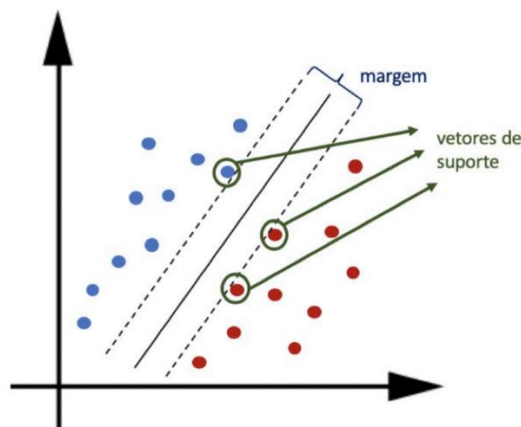


Figura 3: Algoritmo SVM: maximização da margem de separação entre classes através de vetores de suporte e um hiperplano.

Fonte: (Escovedo e Koshiyama, 2020).

O método pode ser aplicado tanto a conjuntos de dados linearmente separáveis quanto a dados não linearmente separáveis. Neste último caso, utiliza-se o chamado *kernel trick*, no qual os dados de entrada são implicitamente mapeados para um espaço de maior dimensionalidade, permitindo que relações não lineares no espaço original sejam tratadas como linearmente separáveis nesse novo espaço (Jakkula, 2006).

Na formulação de regressão, conhecida como *Support Vector Regression* (SVR), o objetivo não é minimizar diretamente o erro quadrático, mas encontrar uma função que apresente o melhor compromisso entre ajuste aos dados e simplicidade do modelo. Para isso, a SVR introduz a função de perda ε -insensitive, na qual desvios menores que um valor ε em torno da função ajustada não são penalizados. Essa abordagem resulta em soluções mais estáveis e menos sensíveis a flutuações locais dos dados experimentais (Smola e Schölkopf, 2004; Smola e Smola, 2002).

A capacidade da SVR de representar comportamentos complexos decorre diretamente do uso de funções *kernel*, que permitem modelar relações altamente não lineares entre as variáveis de entrada e a variável alvo (Vikramaditya Jakkula, 2006). Essa característica é particularmente relevante para a modelagem da resistência mecânica em função da orientação estrutural β , cuja resposta apresenta inversões de tendência e frequentemente assume a forma de curvas em “U”, típicas de rochas anisotrópicas.

Essas propriedades tornam a SVR adequada para problemas de interpolação envolvendo bases de dados reduzidas, como as observadas na resistência anisotrópica de rochas sob diferentes estados de tensões. O controle explícito da complexidade do modelo favorece a obtenção de ajustes suaves, mesmo em contextos com elevada dispersão experimental. Dessa forma, a formulação da SVR conduz a funções interpoladoras contínuas, adequadas para representar a variação da resistência entre pontos experimentais sem a imposição de uma forma funcional pré-definida (Smola e Schölkopf, 2004).

2.4.4 Comparações entre os três métodos

O *Random Forest* responde adequadamente à variabilidade experimental, uma vez que sua estrutura em ensemble reduz a influência de flutuações aleatórias presentes nos dados. Entretanto, sua saída permanece essencialmente determinística do ponto de vista probabilístico, baseada em estimativas pontuais. A *Support Vector Regression* (SVR), por sua vez, destaca-se

pelo controle explícito da complexidade do modelo, favorecendo a generalização em bases de dados reduzidas, embora sua aplicação dependa de escolhas criteriosas de *kernel* e hiperparâmetros. Já a *Gaussian Process Regression* (GPR) diferencia-se por incorporar explicitamente a quantificação da incerteza às previsões, aspecto fundamental em problemas nos quais a confiabilidade da estimativa é tão relevante quanto o valor médio previsto.

Considerando que cada método apresenta respostas distintas aos desafios impostos pela variabilidade experimental, pela não linearidade das relações envolvidas e pelo tratamento da incerteza, a análise comparativa permite explorar suas complementaridades e fundamenta a estratégia metodológica adotada neste estudo.

2.5 Inteligência Artificial Aplicada à Resistência de Rochas

2.5.1 Histórico da aplicação de IA na Mecânica das Rochas

Os métodos empíricos tradicionais são fundamentais para estudos e análises estruturais de engenharia em maciços rochosos. Todavia, a determinação experimental relacionada a esses modelos, via de regra, possui limitações provenientes da alta variabilidade natural das rochas e da dificuldade técnica em mensurar certas propriedades mecânicas. Além da necessidade de correlacionar múltiplas variáveis, tais métodos tornam-se onerosos e demandam ensaios extensos, o que pode inviabilizar sua aplicação em determinadas circunstâncias.

Modelos computacionais de aprendizado de máquina vêm sendo utilizados para modelar sistemas complexos e não lineares em diversas áreas. Atualmente, a aplicação em problemas de engenharia de rochas tem obtido êxito ao suprir as lacunas dos métodos empíricos. Singh et al. (2001), por exemplo, investigaram a viabilidade de prever a resistência à compressão uniaxial (UCS) e o índice de carga pontual (PLT) a partir de bases multivariadas, incluindo propriedades físicas, mineralógicas e a orientação estrutural das rochas por meio de redes neurais artificiais (RNA).

As rochas anisotrópicas apresentam uma relação não linear com múltiplas variáveis simultâneas, o que dificulta a modelagem analítica convencional. Nesse cenário, a utilização de aprendizado supervisionado visa reduzir a dependência de ensaios laboratoriais extensos, fornecendo uma ferramenta alternativa para a estimativa de parâmetros mecânicos essenciais ao projeto de engenharia.

2.5.2 Variáveis normalmente empregadas

A resistência mecânica das rochas é fortemente dependente do estado de tensões aplicado, especialmente do nível de confinamento. Em ensaios de laboratório, a compressão triaxial em corpos de prova cilíndricos é o procedimento padrão para estimar a resistência de pico sob diferentes tensões confinantes (σ_3). Durante o ensaio, o material é submetido a três tensões principais, sendo as tensões de confinamento aqui consideradas iguais ($\sigma_2 = \sigma_3$), enquanto a carga vertical representa a tensão principal maior (σ_1). As variáveis σ_1 e σ_3 são fundamentais para o modelo, pois constituem a informação mínima necessária para qualquer interpolação da resistência em diferentes níveis de carregamento.

Em rochas anisotrópicas, a resistência varia conforme a orientação da estrutura em relação ao campo de tensões, tornando a variável angular β essencial para a capacidade preditiva do modelo. Sem a inclusão de β , o algoritmo seria incapaz de capturar a natureza direcional da rocha. Investigações laboratoriais extensas (Ramamurthy, Rao e Singh, 1993a) demonstram que a resistência de pico atinge seus valores máximos nas orientações de $\beta = 0^\circ$ e 90° . Essa variação gera o formato típico das curvas em "U", que reflete o tipo de anisotropia presente e serve como o principal indicador da fraqueza estrutural do material.

Além das variáveis de carregamento e orientação, parâmetros derivados podem ser empregados para sintetizar o comportamento do material. O grau de anisotropia (A), definido como a razão da resistência máxima para uma determinada tensão confinante, funciona como uma síntese da resposta direcional. A inclusão de A visa reduzir a dimensionalidade do problema e facilitar o aprendizado em bases de dados pequenas, incorporando influências da petrofábrica e da microestrutura que definem o comportamento mecânico (Ramamurthy, Venkatappa Rao, G. e Singh, 1993).

Por fim, parâmetros globais de resistência, amplamente utilizados na prática de engenharia, são incorporados como variáveis auxiliares. Embora Singh et al. (2001) e Cerrato (2024) observem que parâmetros como a resistência à compressão uniaxial da rocha intacta (σ_{ci}) e a constante de material (m_i) possuam significado físico limitado para explicar a causa da anisotropia, sua inclusão é justificada pela ampla disponibilidade e compatibilidade com bancos de dados existentes. Esses parâmetros funcionam como indicadores do comportamento da matriz rochosa, carregando informação útil que auxilia na robustez das previsões do aprendizado de máquina.

3 METODOLOGIA

A metodologia foi estruturada em cinco etapas principais. Inicialmente, realizou-se a compilação, organização e análise exploratória do banco de dados. Em seguida, foram definidas as variáveis de entrada e saída e procedeu-se ao pré-processamento dos dados, incluindo a padronização das escalas. Na terceira etapa, realizou-se a divisão dos dados em conjunto de treinamento e teste, aplicados aos algoritmos *Random Forest* (RF), *Support Vector Regression* (SVR) e *Gaussian Process Regression* (GPR), implementados em linguagem *Python*.

Posteriormente, os modelos foram analisados sob as perspectivas estatística e física, por meio de métricas de desempenho (R^2 , RMSE e MAE), análise de resíduos e verificação da coerência mecânica das curvas preditas em função da orientação e da tensão confinante. Por fim, realizou-se a síntese comparativa dos resultados. Observa-se na Tabela 1 e Figura 4 a estrutura sistemática deste trabalho resumida.

Tabela 1: Resumo de etapas, objetivos e técnicas aplicadas.

Etapas	Objetivos	Técnicas aplicadas
1ª Etapa	Compilação e organização do banco de dados experimental; Análise exploratória inicial dos dados; Verificação da consistência e representatividade do banco.	Compilação de dados de ensaios triaxiais em rochas anisotrópicas a partir da literatura (Cerrato, 2024); Organização do banco em ambiente <i>Python</i> (Pandas); Estatística descritiva global e por nível de σ_3; Histogramas e análise de dispersão; Análise de correlação entre variáveis.
2ª Etapa	Definição das variáveis de entrada e saída do modelo; Incorporação dos efeitos de anisotropia e confinamento; Preparação dos dados para modelagem.	Definição das variáveis independentes (β, σ_3, r, σ_{ci}); Cálculo do ângulo efetivo β_{eff} (período de 180°); Criação de variáveis auxiliares ($\sigma_1 - \sigma_3$ e σ_1/σ_3) para análise física; Padronização das variáveis (StandardScaler); Separação do banco de dados em conjuntos de treino (80%) e teste (20%), com estratificação por σ_3.
3ª Etapa	Desenvolvimento de modelos preditivos para a resistência σ_1; Avaliação do desempenho estatístico dos modelos; Análise da capacidade de generalização.	<i>Random Forest Regressor</i> (RF); <i>Gaussian Process Regression</i> (GPR); <i>Support Vector Regression</i> (SVR); Ajuste de hiperparâmetros (<i>Grid Search</i> e otimização do <i>kernel</i>); Avaliação por métricas R^2, RMSE e MAE; Validação cruzada <i>K-fold</i> (5-fold).
4ª Etapa	Avaliação física dos modelos preditivos; Verificação da coerência mecânica das curvas geradas; Análise do comportamento anisotrópico da resistência.	Geração de curvas $(\sigma_1 - \sigma_3) \times \beta$ para cada nível de σ_3; Representações polares ($0^\circ-360^\circ$) considerando periodicidade de 180°; Verificação da simetria das curvas (função par); Comparação entre valores experimentais e preditos; Análise de resíduos (<i>QQ-plot</i>, histogramas e resíduos vs predição).
5ª Etapa	Comparação final entre os modelos desenvolvidos; Consolidação dos resultados obtidos; Discussão das limitações e implicações do estudo.	Comparação direta do desempenho estatístico dos modelos; Análise gráfica real vs predito (treino e teste); Síntese dos resultados em tabelas e figuras; Discussão metodológica e física dos modelos; Identificação de limitações e perspectivas futuras.

Fonte: (Autora, 2026).

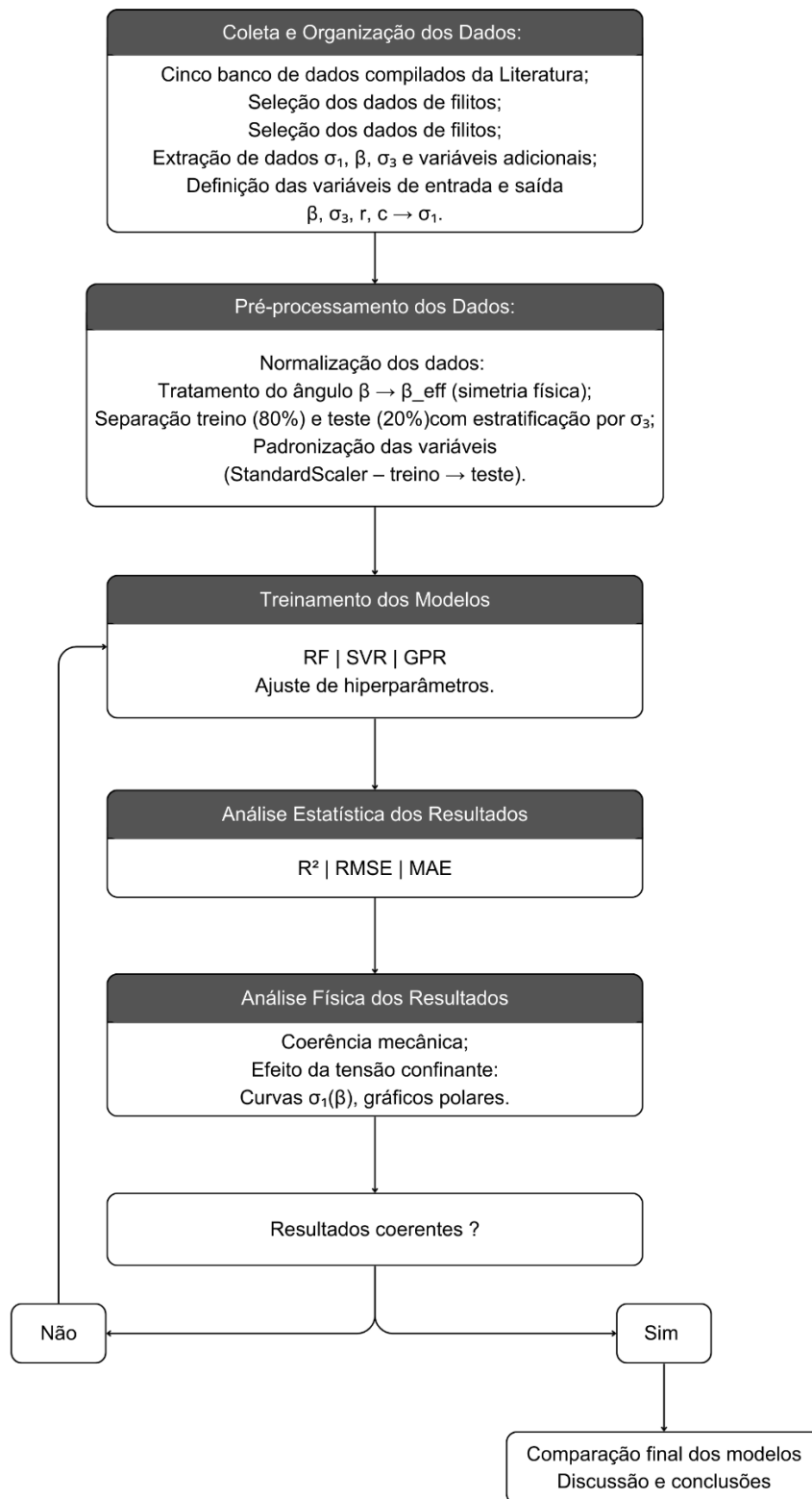


Figura 4: Fluxograma das etapas da metodologia do presente trabalho.

Fonte: (Autora, 2026).

3.1 Banco de dados

O banco de dados utilizado provém de uma compilação realizada por Cerrato (2024), que reúne resultados experimentais de ensaios de compressão triaxial em rochas transversalmente isotrópicas documentados na literatura. Embora o banco original seja heterogêneo e contemple diversas litologias, este estudo realizou um recorte específico para rochas do tipo filito, visando garantir a consistência física das análises.

3.1.1 Origem e Natureza dos Dados

Os dados consistem em resultados laboratoriais que descrevem a resistência de pico (σ_1) sob diferentes níveis de tensão confinante (σ_3). Após a seleção, o banco de dados final totalizou 161 corpos de prova, subdivididos em 5 variedades distintas de filitos (Tabela 2). As nomenclaturas das variáveis originais e derivadas estão detalhadas na Tabela 3, e a relação completa dos dados brutos consta no Apêndice deste trabalho.

Tabela 2: Lista dos filitos que basearam o banco de dados, incluindo referência bibliográfica.

Rocha	Referência
<i>Carbonaceous Phyllite</i>	(Singh J. 1988)
<i>Quartzite Phyllite</i>	(Singh J. 1988)
<i>Micaceous Phyllite</i>	(Singh J. 1988)
<i>Phyllites dry</i>	(Kumar, 2006)
<i>Moretown Phyllite</i>	(Donath, 1972)

Fonte: (Autora, 2026).

Tabela 3: Nomenclatura das variáveis do banco de dados.

Nomenclatura	Variável
σ_1	Resistência de pico
σ_3	Tensão confinante
β	Ângulo de orientação entre o plano de anisotropia e a direção da tensão principal maior (σ_1)
A	Grau de anisotropia
σ_{ci}	Resistência à compressão uniaxial
c	Coesão
φ	Ângulo de atrito

Fonte: (Autora, 2026).

3.1.2 Variáveis de Entrada e Saída

A modelagem foi estruturada separando as informações em variáveis de saída (alvo) e variáveis de entrada (*features*), selecionadas com base na interpretação física da anisotropia:

- Variável de Saída: resistência de pico (σ_1).
- Variáveis de Entrada: tensão confinante (σ_3), ângulo de orientação da foliação (β), grau de anisotropia (A) e resistência à compressão uniaxial (σ_{ci}).

3.1.3 Processamento e Variáveis Derivadas

A partir dos dados básicos, foram incorporados parâmetros calculados por Cerrato (2024), como a coesão (c) e o ângulo de atrito (φ). Adicionalmente, foram geradas variáveis derivadas como a diferença de tensões ($\sigma_1 - \sigma_3$) e a razão de tensões (σ_1/σ_3), utilizadas exclusivamente como auxiliares para a validação da consistência física dos modelos, não integrando o conjunto de treinamento direto.

A variável angular β foi tratada e convertida para β_{eff} , introduzindo uma periodicidade de 180° que assegura o respeito à simetria física esperada em rochas foliadas.

3.1.4 Origem e Caracterização dos Dados de Filito

A base de dados segmentada é composta por resultados de ensaios de compressão triaxial realizados em cinco variedades de filitos, compilados por Cerrato (2024) a partir de três estudos de referência na literatura:

- Singh J. (1988): Dados de Filitos Quartzítico, Carbonoso e Micáceo extraídos da região do Himalaia. As amostras apresentam foliação marcante e veios de quartzo concordantes, resultando em um comportamento mecânico dependente da orientação estrutural.
- Donath (1972): Ensaios no Filito *Moretown* (*Vermont, EUA*), capturando uma ampla gama de modos de deformação e deslizamento ao longo da clivagem em diversas orientações.
- Kumar (2006): Filito, obtido na margem direita do rio *Bhagirathi*, em *Joshiyara* (próximo a *Uttarkashi*), composto principalmente por micas e minerais quartzo-feldspáticos.

3.2 Pré-processamento dos Dados

Antes do treinamento dos modelos, realizou-se o pré-processamento dos dados com o objetivo de garantir coerência física, estabilidade numérica e consistência estatística para o treinamento e avaliação dos algoritmos de aprendizado de máquina.

O primeiro passo consistiu no tratamento da variável angular β , associada à orientação dos planos de anisotropia. Visto que a resistência mecânica de rochas foliadas apresenta simetria em relação aos planos de foliação, dependendo da orientação relativa entre estrutura e carregamento, e não do sinal do ângulo, a variável β foi transformada em um ângulo efetivo (β_{eff}). Esta transformação impõe ao modelo a simetria física esperada, assegurando que as curvas $\sigma_1 - \sigma_3$ em função de β apresentem o comportamento par característico.

Dando continuidade ao processo, as variáveis de entrada e saída, previamente estabelecidas, foram organizadas para a estruturação dos conjuntos de dados. O banco de dados foi então dividido em conjuntos de treinamento (80%) e teste (20%), de modo a permitir a avaliação da capacidade de generalização dos modelos em dados não utilizados no ajuste. Essa proporção foi adotada em função do tamanho reduzido do banco de dados, de modo a priorizar maior disponibilidade de amostras para o aprendizado dos padrões físicos.

Para assegurar a reprodutibilidade da divisão, utilizou-se o parâmetro *random_state* igual a 42. Além disso, aplicou-se a estratificação das amostras com base na tensão de confinamento (σ_3), garantindo que os diferentes níveis de confinamento apareçam proporcionalmente em ambos os conjuntos (treino/teste), evitando, assim, a introdução de vieses no aprendizado.

Não foram identificados valores ausentes ou inconsistentes no banco de dados após a etapa de filtragem por litotipo. Também não foi necessária a remoção de outliers, visto que os valores extremos correspondem a comportamentos físicos documentados na literatura.

Por fim, as variáveis de entrada foram padronizadas por meio do método *StandardScaler* (Scikit-learn, 2026), que transforma os dados para média zero e desvio padrão unitário. O ajuste do escalonador foi realizado exclusivamente no conjunto de treinamento e posteriormente aplicado ao conjunto de teste, prática essencial para evitar o vazamento de informação (*data leakage*). Esta etapa assegura que variáveis com diferentes ordens de grandeza não exerçam influência desproporcional no treinamento dos modelos.

3.3 Treinamento dos Modelos

Os modelos preditivos foram implementados em ambiente *Python* utilizando a biblioteca *scikit-learn*, amplamente empregada em aplicações de aprendizado de máquina supervisionado. O treinamento foi realizado separadamente para cada algoritmo respeitando as particularidades metodológicas de cada técnica.

Em todos os casos, o treinamento foi conduzido exclusivamente sobre o conjunto de dados de treinamento, conforme definido na etapa anterior, preservando o conjunto de teste para avaliação independente do desempenho e da capacidade de generalização dos modelos.

Os modelos foram treinados individualmente para cada litotipo, evitando a mistura de comportamentos mecânicos distintos e permitindo que cada algoritmo capturasse as particularidades físicas associadas a cada variedade de filito.

3.3.1 *Random Forest*

O modelo *Random Forest* foi treinado com base em um conjunto de árvores de decisão construídas a partir de subconjuntos aleatórios dos dados. Essa estratégia de ensemble reduz a variância do modelo e aumenta sua robustez.

Foram utilizadas 800 árvores individuais ($n_estimators = 800$), número escolhido de modo a garantir a estabilidade das previsões. Não foi imposto limite máximo à profundidade das árvores ($max_depth = None$), permitindo que o modelo capturasse relações não lineares complexas entre as variáveis de entrada e a resistência de pico. Os parâmetros mínimos para divisão dos nós foram mantidos nos valores padrão do *scikit-learn*, assegurando flexibilidade na construção das árvores.

O parâmetro $max_features = "sqrt"$ foi adotado para definir o número de variáveis consideradas em cada divisão, promovendo diversidade entre as árvores. O uso de amostragem com reposição ($bootstrap = True$) permitiu que cada árvore fosse treinada a partir de uma subamostra aleatória dos dados. A reprodutibilidade do treinamento foi garantida por meio da fixação do parâmetro $random_state$.

A avaliação do desempenho e da capacidade de generalização do modelo foi realizada por meio de validação cruzada do tipo *K-fold* com cinco partições.

3.3.2 *Gaussian Process Regression*

O treinamento do modelo de Regressão por Processos Gaussianos foi realizado a partir da definição explícita de uma função de covariância (*kernel*), responsável por controlar a suavidade, a correlação espacial e o tratamento do ruído nos dados.

Foi adotado um *kernel* do tipo *Radial Basis Function* (RBF), com escalas de comprimento independentes para cada variável de entrada, permitindo que o modelo atribuísse diferentes relevâncias às variáveis físicas envolvidas.

O modelo foi configurado com normalização da variável de saída (*normalize_y = True*), etapa essencial para garantir a estabilidade numérica e a convergência do processo de otimização. A calibração dos hiperparâmetros do kernel foi realizada por meio de múltiplas reinicializações do otimizador (*n_restarts_optimizer = 10*), reduzindo a dependência de condições iniciais.

A avaliação do desempenho do modelo foi conduzida por validação cruzada *K-fold* com cinco partições, permitindo analisar a consistência das previsões ao longo das diferentes subdivisões do conjunto de dados.

3.3.3 *Support Vector Regression*

O modelo de *Support Vector Regression* foi treinado utilizando *kernel* do tipo *Radial Basis Function* (RBF), adequado para a representação de relações não lineares suaves entre as variáveis de entrada e a resistência de pico.

A seleção dos hiperparâmetros foi realizada por meio de uma busca em grade (*Grid Search*), considerando os parâmetros de regularização (C), largura da margem insensível (ϵ) e influência dos pontos de treinamento (γ). Essa estratégia permitiu identificar a combinação de parâmetros que maximiza a capacidade de generalização do modelo, evitando ajustes excessivamente rígidos ou excessivamente flexíveis.

A validação dos hiperparâmetros foi conduzida por validação cruzada *K-fold* com cinco partições, utilizando o coeficiente de determinação (R^2) como métrica principal de desempenho. O modelo final foi então treinado com os parâmetros ótimos obtidos e avaliado posteriormente no conjunto de teste.

3.4 Avaliação do Desempenho dos Modelos

3.4.1 Avaliação estatística

O desempenho dos modelos preditivos foi avaliado por meio de métricas estatísticas que quantificam a qualidade das estimativas da resistência de pico (σ_1) em função da orientação estrutural (β), da tensão confinante (σ_3) e dos parâmetros do material (R e σ_{ci}).

A principal métrica utilizada foi o coeficiente de determinação (R^2), que expressa a proporção da variabilidade da resistência σ_1 explicada pelas variáveis de entrada do modelo. Valores de R^2 próximos de 1 indicam excelente ajuste entre os valores observados e preditos, enquanto valores próximos de zero indicam baixo poder explicativo. Valores negativos indicam desempenho inferior ao de uma estimativa baseada apenas na média dos dados.

Essa métrica foi empregada para comparar o desempenho nos conjuntos de treinamento e teste, permitindo a identificação de possíveis efeitos de sobreajuste (*overfitting*), bem como para avaliar a estabilidade dos modelos por meio da validação cruzada. A definição matemática do coeficiente de determinação é apresentada na Equação (1).

$$R^2 = 1 - \frac{\sum(\sigma_{1,real} - \sigma_{1,pred})^2}{\sum(\sigma_{1,real} - \bar{\sigma}_1)^2} \quad (1)$$

Complementarmente, foram utilizadas métricas baseadas em erro absoluto, com interpretação física direta:

- *Root Mean Squared Error* (RMSE): mede o erro quadrático médio entre os valores observados e preditos, penalizando de forma mais severa erros de maior magnitude. Do ponto de vista físico, o RMSE representa o erro típico do modelo ao estimar a resistência máxima sob compressão, sendo expresso em unidades de tensão (MPa). Sua formulação matemática é apresentada na Equação (2).

$$RMSE = \sqrt{\frac{1}{n} \sum (\sigma_{1,real} - \sigma_{1,pred})^2} \quad (2)$$

- *Mean Absolute Error* (MAE): mede o erro médio absoluto, sem penalização adicional para valores extremos. Essa métrica fornece uma estimativa direta do erro médio esperado ao utilizar o modelo para prever σ_1 em um ensaio com

orientação β e tensão confinante σ_3 conhecidas. Sua formulação matemática é apresentada na Equação (3).

$$MAE = \frac{1}{n} \sum |\sigma_{1,real} - \sigma_{1,pred}| \quad (3)$$

A utilização conjunta dessas métricas permite avaliar tanto a capacidade explicativa global do modelo (R^2) quanto a magnitude física dos erros associados às previsões (RMSE e MAE), fornecendo uma análise estatística mais robusta.

3.4.2 Estatísticas Descritivas do Banco de Dados

Além das métricas de desempenho dos modelos, foram calculadas estatísticas descritivas com o objetivo de caracterizar o banco de dados experimental utilizado na modelagem. Essas métricas não avaliam diretamente os modelos, mas fornecem informações importantes sobre a distribuição e a variabilidade dos dados.

Foram analisadas medidas de tendência central, como média, mediana e moda, utilizadas para identificar o grau de simetria da distribuição dos dados. O desvio padrão e a variância foram empregados para quantificar a dispersão dos valores de resistência, refletindo a heterogeneidade experimental típica de rochas anisotrópicas.

O coeficiente de variação foi utilizado como medida adimensional de dispersão, permitindo a comparação da variabilidade entre variáveis com escalas distintas. Adicionalmente, os coeficientes de assimetria e curtose foram analisados para identificar possíveis vieses experimentais ou concentrações de dados em determinadas faixas de orientação β ou níveis de confinamento.

3.4.3 Métricas de Validação e Diagnóstico

A validação e o diagnóstico dos modelos foram realizados por meio da análise dos resíduos, definidos como a diferença entre a resistência observada experimentalmente e a resistência estimada pelo modelo.

Foram utilizados gráficos de resíduos e *QQ-plots*, que permitem avaliar se os erros apresentam comportamento aleatório ou se há padrões sistemáticos associados às previsões. A

presença de resíduos aproximadamente simétricos e sem tendência indica bom ajuste estatístico do modelo.

Adicionalmente, foi aplicada validação cruzada do tipo *K-fold*, sendo analisados o valor médio do coeficiente de determinação (R^2) e seu desvio padrão entre as partições. Essa abordagem demonstra que o desempenho dos modelos não depende de uma única divisão específica dos dados, reforçando a robustez e a capacidade de generalização das previsões.

3.4.4 Avaliação física e coerência mecânica

A coerência física e mecânica dos modelos foi avaliada por meio da análise gráfica do comportamento da resistência em função da orientação estrutural e do nível de confinamento. Para esse fim, foram geradas curvas de $(\sigma_1 - \sigma_3)$ em função de β , permitindo avaliar a variação da tensão desviadora e os mecanismos de falha característicos de rochas anisotrópicas.

A utilização da diferença $(\sigma_1 - \sigma_3)$ permite reduzir o efeito direto do confinamento, possibilitando a comparação do grau de anisotropia sob diferentes níveis de σ_3 . A análise das curvas individualizadas para cada tensão confinante permitiu observar que o mecanismo de ruptura e a intensidade da anisotropia podem se modificar com o aumento do confinamento, resultando em intensificação ou suavização do comportamento direcional.

Adicionalmente, foram geradas curvas da razão σ_1/σ_3 em função de β , bem como seus respectivos diagramas polares, com o objetivo de avaliar a anisotropia da resistência e a capacidade dos modelos em reproduzir a simetria física esperada. Em rochas foliadas, a resistência mecânica depende da orientação relativa entre o plano de anisotropia e a direção da tensão principal maior, sendo independente do sentido do carregamento, característica que deve ser refletida nas curvas preditas.

Também foram elaborados gráficos de dispersão entre valores preditos e observados de σ_1 , tanto para os conjuntos de treinamento quanto de teste, permitindo avaliar visualmente a qualidade do ajuste e possíveis tendências sistemáticas.

3.4.5 Comparação entre os Modelos

A comparação final entre os modelos foi realizada com base nas métricas estatísticas e nas análises físicas previamente apresentadas. Foram construídos histogramas para

comparação dos valores de R^2 e RMSE dos três algoritmos, além de gráficos de dispersão entre valores reais e preditos.

Adicionalmente, foi elaborada uma tabela síntese contendo as métricas globais de desempenho, permitindo uma comparação direta entre os modelos quanto à precisão, à robustez e à coerência física das previsões.

4 RESULTADOS E DISCUSSÕES

Neste capítulo são apresentados e analisados os resultados obtidos a partir da modelagem da resistência de pico dos filitos anisotrópicos. A apresentação dos resultados é realizada por meio de tabelas e gráficos, de forma a facilitar a interpretação do comportamento mecânico dos diferentes litotipos e a avaliação dos modelos preditivos aplicados. O código desenvolvido para a implementação das etapas metodológicas descritas neste estudo encontra-se integralmente apresentado no Apêndice 1, para fins de reprodutibilidade e transparência da pesquisa.

4.1 Apresentação dos bancos de dados e caracterização geral

Os resultados baseiam-se em cinco bancos de dados experimentais de filitos, extraídos da literatura (Cerrato, 2024), os quais apresentam diferenças quanto ao número de ensaios, aos níveis de tensão confinante e à magnitude da resistência de pico (σ_1).

Essas diferenças conferem aos bancos de dados distintos graus de anisotropia e variabilidade mecânica, justificando tanto a análise individual de cada litotipo quanto a comparação entre eles ao longo deste capítulo.

Observam-se na Tabela 4 as principais características dos bancos de dados utilizados, servindo de referência para a interpretação dos resultados apresentados nas seções seguintes.

Tabela 4: Resumo das principais características dos bancos de dados utilizados.

Litotipo	Nº de ensaios	Níveis de β (°)	Níveis de σ_3 (MPa)	σ_1 mínimo – máximo (MPa)
<i>Micaceous Phyllite</i>	35	0, 15, 30, 45, 60, 75, 90	5, 15, 30, 50, 70	62,6 – 335,4
<i>Quartzite Phyllite</i>	35	0, 15, 30, 45, 60, 75, 90	5, 15, 30, 50, 70	58,8 – 338,2
<i>Phyllites dry</i>	35	0, 15, 30, 45, 60, 75, 90	0, 5, 15, 30, 60	30,3 – 357,6
<i>Carbonaceous Phyllite</i>	35	0, 15, 30, 45, 60, 75, 90	5, 15, 30, 50, 70	32,7 – 326,8
<i>Moretown Phyllite</i>	21	0, 15, 30, 45, 60, 75, 90	50, 100, 200	135,9 – 738,3

Fonte: (Autora, 2026).

4.2 Análise exploratória do comportamento anisotrópico

A resistência mecânica dos filitos apresenta forte dependência direcional em função da orientação da foliação em relação à direção da tensão principal maior. A análise exploratória

dos bancos de dados permitiu avaliar essa anisotropia de forma direta, considerando a variação da resistência de pico (σ_1) com o ângulo β e com a tensão confinante (σ_3).

Os resultados indicam que a resistência de pico dos filitos não permanece constante ao longo das diferentes orientações da foliação e níveis de tensão confinante, evidenciando um comportamento anisotrópico típico e apresentando padrões consistentes e reproduzíveis para os diferentes litotipos analisados.

Para todos os litotipos analisados, observa-se uma variação sistemática de σ_1 com o ângulo β . De modo geral, os valores máximos de σ_1 tendem a ocorrer nas orientações próximas a 0° e aproximadamente 90° , enquanto os valores mínimos concentram-se em ângulos intermediários. Para o filito *Carbonaceous*, os mínimos ocorrem aproximadamente entre 40° e 75° ; para o filito *Moretown*, entre 30° e 60° ; para o filito *Dry*, entre 35° e 75° ; para o filito *Quartzite*, entre 30° e 75° ; e para o filito *Micaceous*, entre 30° e 65° conforme apresentado na Figura 5. Esses resultados são compatíveis com padrões observados em estudos anteriores sobre rochas anisotrópicas (Ramamurthy, Venkatappa Rao e Singh, 1993b; Ramamurthy, Venkatappa Rao e Singh J., 1988).

Os gráficos de σ_1 em função de β_{eff} apresentam padrões suaves e contínuos, reforçando a existência de uma relação não linear entre a resistência de pico e a orientação estrutural, conforme ilustrado na Figura 5. Essa variação é expressa por curvas de resistência em função da orientação β , que tipicamente assumem um formato em “U”, característico de materiais anisotrópicos (Ramamurthy, Venkatappa Rao e Singh J., 1988).

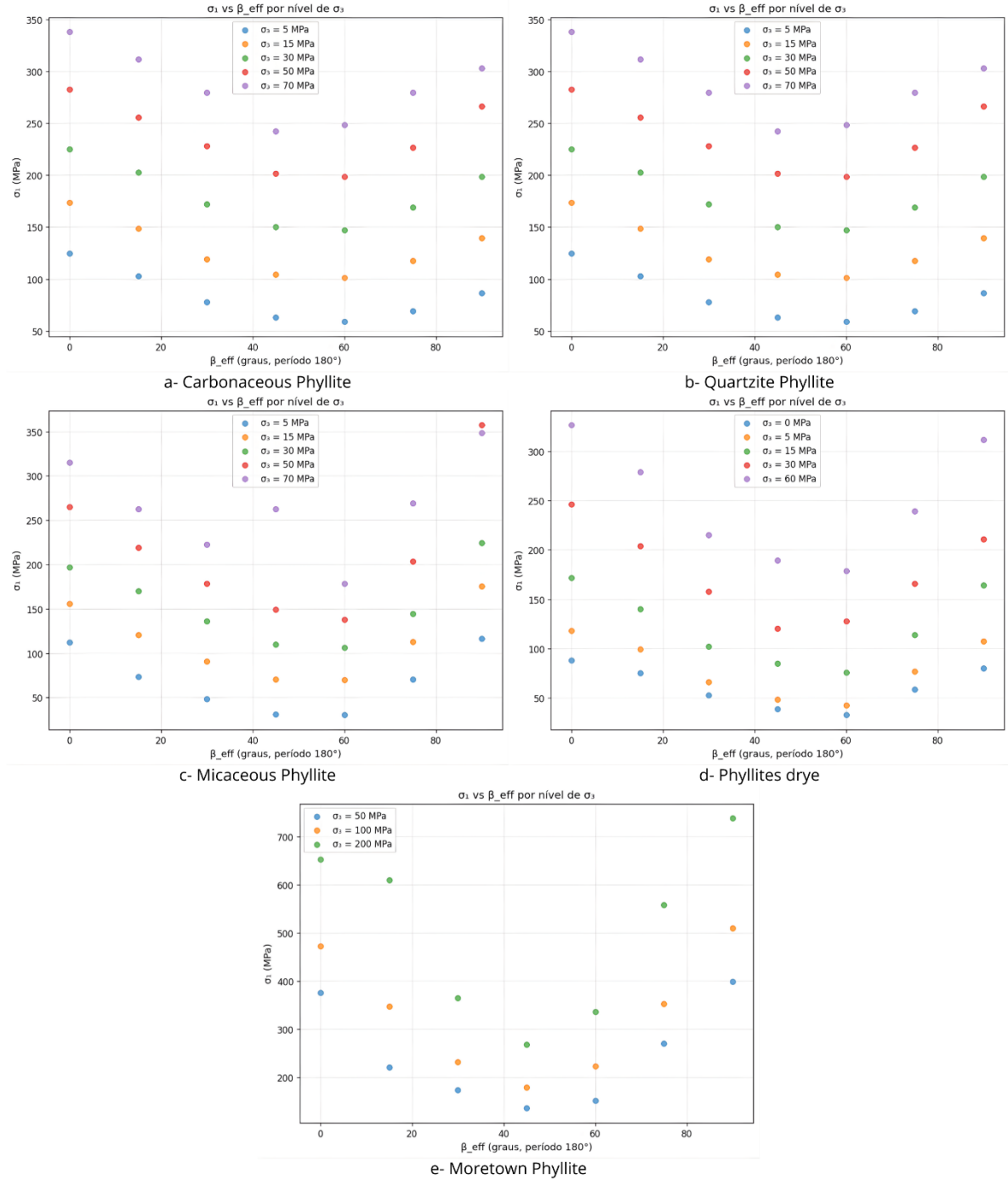


Figura 5: Gráficos de σ_1 vs β_{eff} para (a) *Carbonaceous Phyllite*, (b) *Quartzite Phyllite*, (c) *Micaceous Phyllite*, (d) *Phyllites drye* e (e) *Moretown Phyllite*.

Fonte: (Autora, 2026).

A análise dos dados separados por níveis de tensão confinante, também apresentada na Figura 5, evidencia que o aumento de σ_3 promove um aumento global da resistência de pico em todas as orientações da foliação. Observa-se, adicionalmente, que o grau de anisotropia relativo tende a diminuir à medida que σ_3 aumenta.

Esse comportamento sugere que, à medida que a pressão confinante aumenta, ocorre o fechamento progressivo das descontinuidades e a mobilização do atrito, reduzindo o contraste relativo associado à anisotropia, sem eliminar completamente o controle estrutural (Ramamurthy, Rao e Singh, 1993a; Singh et al., 2015).

Com o objetivo de quantificar o grau de anisotropia de forma relativa, foi adotado um índice definido pela razão entre os valores máximo e mínimo da resistência de pico para cada nível de tensão confinante. A variação desse índice com σ_3 , apresentada na Figura 6, evidencia que o confinamento exerce papel fundamental na anisotropia da resistência. Para baixos valores de σ_3 , a variação de σ_1 com β é mais acentuada, indicando maior influência da orientação da foliação no processo de ruptura.

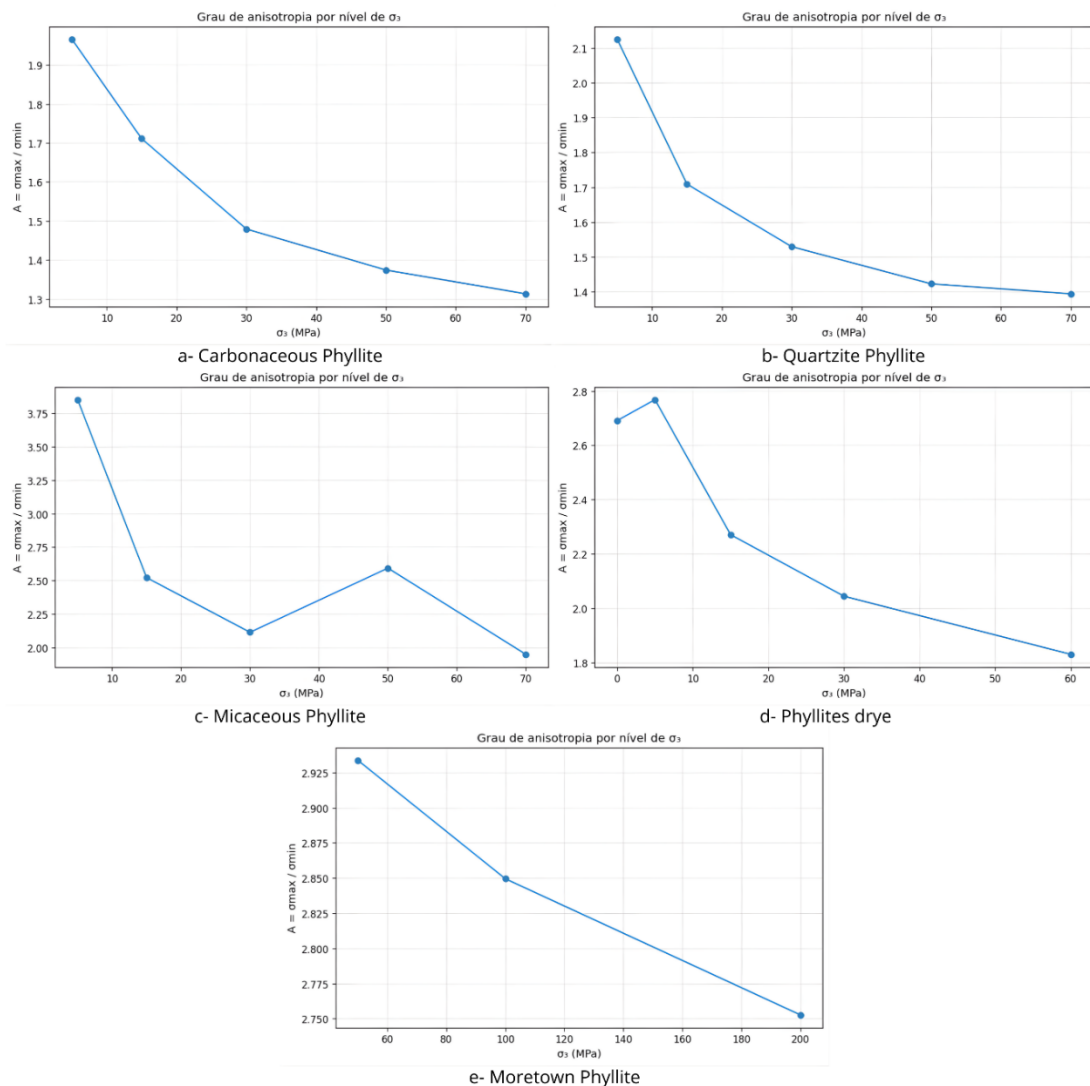


Figura 6: Gráfico do grau de anisotropia por nível de σ_3 , analisados para (a) *Carbonaceous Phyllite*, (b) *Quartzite Phyllite*, (c) *Micaceous Phyllite*, (d) *Phyllites drye* e (e) *Moretown Phyllite*.

Fonte: (Autora, 2026).

Adicionalmente, a análise da variável $(\sigma_1 - \sigma_3)$ em função de β_{eff} , apresentada na Figura 7, permite isolar o efeito do confinamento, destacando que, embora a anisotropia persista em níveis elevados de σ_3 , sua intensidade relativa tende a diminuir. Esse padrão é consistente com o comportamento esperado de rochas anisotrópicas submetidas a elevados estados de tensão confinante (Ramamurthy, Rao e Singh, 1993).

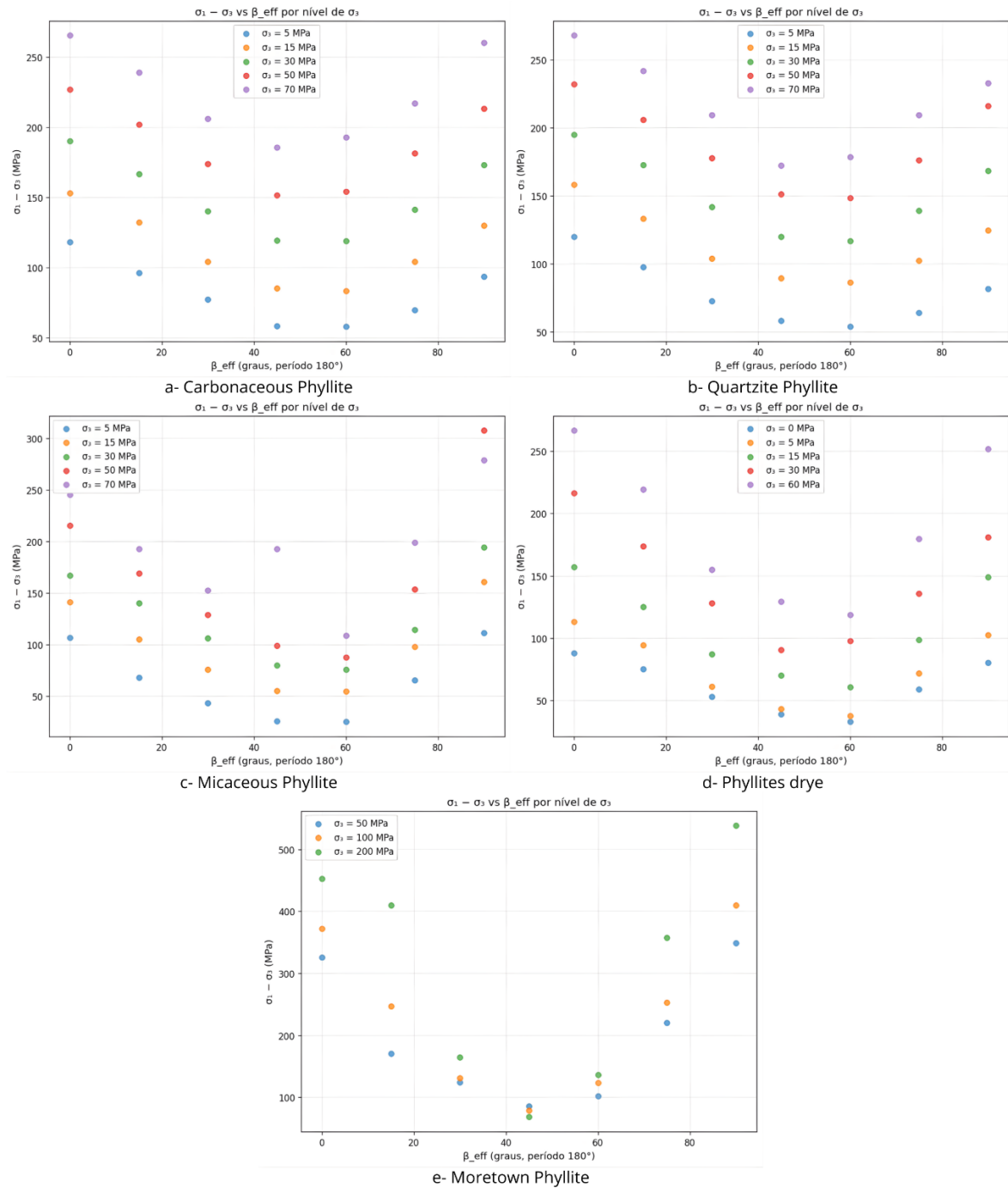


Figura 7: Gráfico de dispersão que mostra a relação entre a diferença de tensões ($\sigma_1 - \sigma_3$) e o ângulo efetivo (β_{eff}) para diferentes níveis de tensão confinante (σ_3)

Fonte: (Autora, 2026).

4.2.1 Comparação do comportamento anisotrópico entre os diferentes filitos

A comparação entre os diferentes litotipos evidencia que a intensidade da anisotropia e a magnitude da resistência de pico variam significativamente entre os bancos de dados analisados, refletindo diferenças mineralógicas e estruturais entre os filitos.

O filito de *Moretown* destaca-se por apresentar os maiores valores de resistência de pico, atingindo valores superiores a 600 MPa, enquanto os demais litotipos raramente ultrapassam 350 MPa (Tabela 4). Além da maior resistência absoluta, esse litotipo mantém contraste direcional significativo mesmo sob elevados níveis de tensão confinante, indicando forte controle estrutural do comportamento mecânico.

Os filitos *Carbonaceous* e *Quartzite* apresentam os menores valores do índice de anisotropia relativo, com valores de A próximos de 1,3 a 1,4 para maiores níveis de σ_3 , sugerindo comportamento mecânico mais homogêneo. Em contrapartida, os filitos *Micaceous* e *Dry* exibem anisotropia mais pronunciada, associada à maior influência da orientação da foliação na resistência de pico (Tabela 5).

Tabela 5: Índice de anisotropia (A) por litotipo e nível de σ_3 .

<i>Carbonaceous</i>		<i>Quartzite</i>		<i>Micaceous</i>		<i>Dry</i>		<i>Moretown</i>	
Níveis de σ_3	Grau de anisotropia relativo (A)	Níveis de σ_3	Grau de anisotropia relativo (A)	Níveis de σ_3	Grau de anisotropia relativo (A)	Níveis de σ_3	Grau de anisotropia relativo (A)	Níveis de σ_3	Grau de anisotropia relativo (A)
5	2,0	5	2,1	5	3,9	0	2,7	50	2,9
15	1,7	15	1,7	15	2,5	5	2,8	100	2,8
30	1,5	30	1,5	30	2,1	15	2,3	200	2,8
50	1,4	50	1,4	50	2,6	30	2,0		
70	1,3	70	1,4	70	1,9	60	1,8		

Fonte: (Autora, 2026).

De forma geral, a análise exploratória dos bancos de dados confirma que a resistência de pico dos filitos é fortemente dependente da orientação da foliação e do nível de tensão confinante, apresentando anisotropia direcional não linear. Embora o aumento do confinamento reduza o contraste relativo associado à anisotropia, o controle estrutural não é completamente eliminado. Esses resultados reforçam a necessidade de modelos capazes de representar de forma contínua e integrada os efeitos da orientação estrutural e do confinamento sobre a resistência mecânica, fornecendo base conceitual para a aplicação dos modelos preditivos apresentados nas seções seguintes.

4.3 Desempenho dos modelos de interpolação da resistência de pico

Nesta seção avalia-se a capacidade dos modelos de aprendizado de máquina em interpolar a resistência de pico (σ_1) dos filitos anisotrópicos, considerando simultaneamente os efeitos da orientação da foliação (β_{eff}), da tensão confinante (σ_3) e das propriedades mecânicas associadas. A análise baseia-se nos padrões físicos identificados na Seção 4.2, permitindo verificar não apenas o desempenho estatístico dos modelos, mas também sua coerência com o comportamento mecânico esperado para rochas anisotrópicas.

4.3.1 Avaliação do modelo *Random Forest*

O modelo *Random Forest* foi aplicado à interpolação da resistência de pico (σ_1) com o objetivo de avaliar sua capacidade de representar o comportamento anisotrópico do banco de dados em estudo. A avaliação do desempenho foi realizada a partir da separação dos dados em conjuntos de treinamento e teste (Tabela 6), permitindo analisar tanto a capacidade de ajuste quanto a generalização do modelo.

Tabela 6: Tamanhos dos conjuntos de treinamento e teste utilizados na avaliação dos modelos.

Litotipo	Tamanhos dos conjuntos de treino	Tamanhos dos conjuntos de teste
<i>Carbonaceous</i>	28	7
<i>Quartzite</i>	28	7
<i>Micaceous</i>	28	7
<i>Dry</i>	28	7
<i>Moretown</i>	16	5

Fonte: (Autora, 2026).

O desempenho do modelo foi avaliado por meio das métricas: coeficientes de determinação (R^2), erro médio absoluto (MAE) e raiz do erro quadrático médio (RMSE), considerando separadamente os conjuntos de treinamento e teste para cada litotipo analisado.

De modo geral, os elevados valores de R^2 obtidos no conjunto de treinamento indicam que o modelo foi capaz de capturar adequadamente as relações entre as variáveis de entrada e a resistência de pico. A comparação entre os valores de R^2 de treinamento e teste permite avaliar a capacidade de generalização do modelo, sendo observada boa concordância entre esses valores para a maioria dos litotipos, o que sugere ausência de sobreajuste significativo, conforme apresentado na Tabela 7.

Tabela 7: Desempenho do modelo *Random Forest* na predição da resistência de pico (σ_1).

Litotipo	R ² (treino)	R ² (teste)	Amplitude de σ_1 (MPa)	MAE (teste)	Erro médio	RMSE (teste)	Erro médio
<i>Carbonaceous</i>	0,991	0,955	272,8	12,403	5%	13,88	5%
<i>Quartzite</i>	0,991	0,966	279,4	11,496	4%	12,297	4%
<i>Micaceous</i>	0,967	0,81	327,3	21,783	7%	24,411	7%
<i>Dry</i>	0,978	0,882	294,1	14,264	5%	19,124	7%
<i>Moretown</i>	0,928	0,443	602,4	110,985	18%	116,178	19%

Fonte: (Autora, 2026).

Para os filitos *Carbonaceous* e *Quartzite*, os valores de R² no conjunto de teste (0,955 e 0,966, respectivamente) indicam excelente capacidade explicativa do modelo. De forma consistente, os valores de RMSE e MAE obtidos para esses litotipos são baixos quando comparados à amplitude total da resistência de pico observada experimentalmente. Para o filito *Carbonaceous*, os valores de RMSE ($\approx 13,9$ MPa) e MAE ($\approx 12,4$ MPa) representam cerca de 5% da variação total de σ_1 , enquanto para o filito *Quartzite* os valores de RMSE ($\approx 12,3$ MPa) e MAE ($\approx 11,5$ MPa) correspondem igualmente a aproximadamente 4% da amplitude observada. Esses resultados evidenciam elevada precisão do modelo na interpolação da resistência de pico para esses materiais.

Os filitos *Micaceous* e *Dry* apresentaram desempenho intermediário. Embora os valores de R² no conjunto de teste (0,81 e 0,882, respectivamente) indiquem boa capacidade de generalização, observa-se aumento dos valores de RMSE e MAE em relação aos litotipos anteriores, conforme apresentado na Tabela 7. Esses resultados refletem a maior variabilidade estrutural desses litotipos, associada à anisotropia mais pronunciada, o que impõe maior complexidade à interpolação da resistência de pico.

Em contraste, o filito de *Moretown* apresentou desempenho inferior no conjunto de teste, com valor de R² de 0,443, indicando dificuldade do modelo em generalizar o comportamento mecânico desse litotipo. Os valores elevados de RMSE (≈ 116 MPa) e MAE (≈ 111 MPa) refletem a ampla faixa de resistência de pico observada experimentalmente (≈ 602 MPa), bem como o número reduzido de ensaios disponíveis. Ainda assim, quando interpretados de forma relativa, esses erros correspondem a aproximadamente 18–19% da amplitude total de σ_1 , indicando que, embora o modelo capture parcialmente a tendência global, sua precisão é limitada para esse banco de dados específico.

Observa-se na Figura 8 a comparação entre os valores experimentais e preditos de σ_1 para os conjuntos de treinamento e teste. Os pontos se distribuem predominantemente em torno

da linha, indicando boa concordância entre os valores observados e estimados pelo modelo. Uma dispersão moderada é esperada, considerando a variabilidade intrínseca dos dados experimentais e o número limitado de ensaios disponíveis para alguns litotipos. Em particular, o filito de *Moretown* apresenta maior dispersão, especialmente no conjunto de teste, o que é consistente com os resultados obtidos pelas métricas quantitativas.

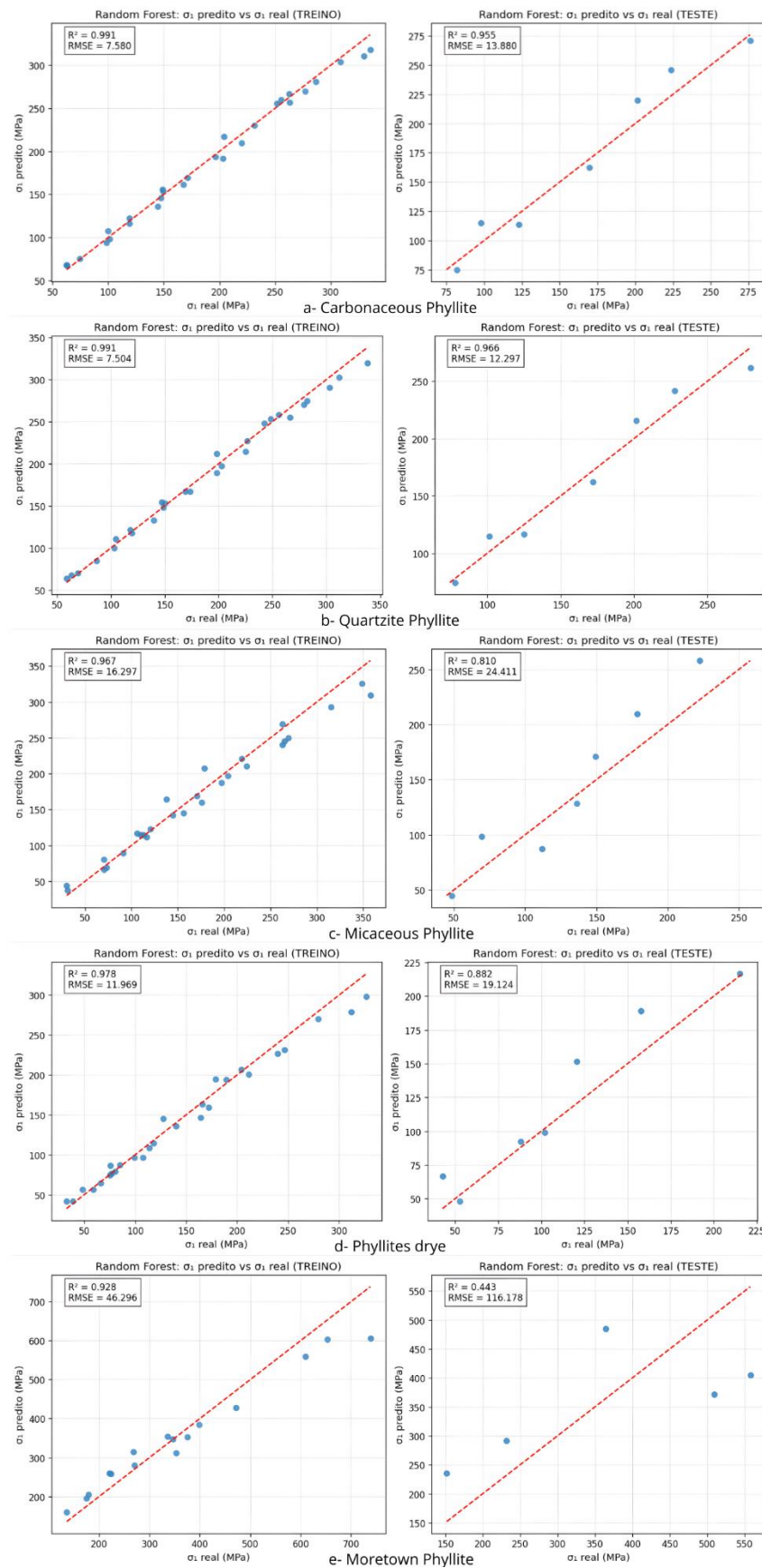


Figura 8: Comparação entre valores reais e preditos de σ_1 pelo modelo *Random Forest*.

Fonte: (Autora, 2026).

As curvas interpoladas de $\sigma_1 - \sigma_3$ em função de β_{eff} obtidas pelo modelo *Random Forest* reproduzem de forma consistente os padrões anisotrópicos identificados na análise exploratória. Conforme ilustrado na Figura 9, as curvas preditas seguem o padrão esperado “U”, com valores mínimos de resistência em ângulos intermediários de β_{eff} e valores máximos próximos a 0° e 90° , além de apresentar aumento global da resistência de pico com o incremento da tensão confinante para todos os litotipos analisados.

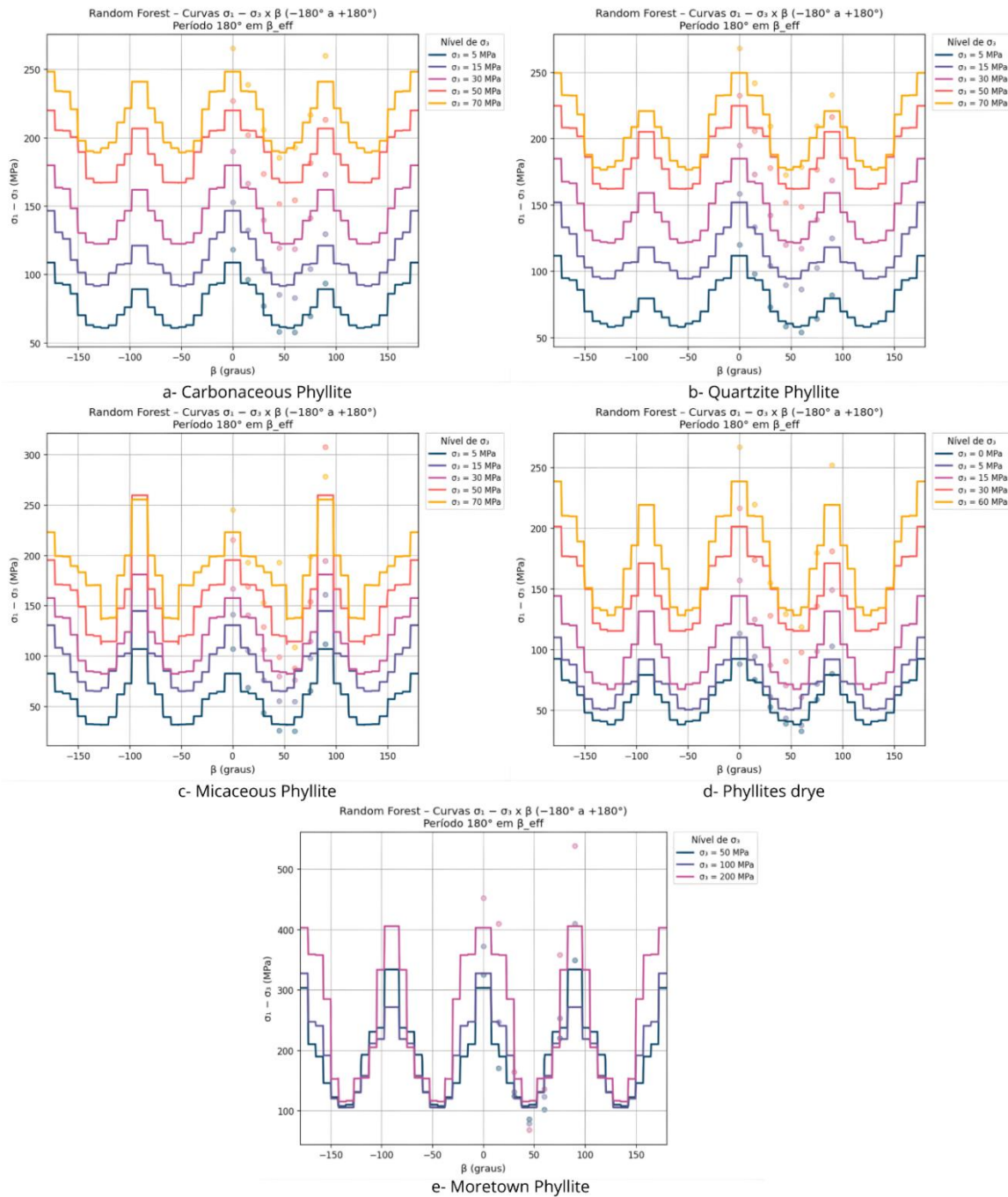


Figura 9: Curvas $\sigma_1 - \sigma_3 \times \beta_{eff}$ preditas pelo *Random Forest*.

Fonte: (Autora, 2026).

Esses resultados indicam que o modelo Random Forest foi capaz de capturar não apenas a relação estatística entre as variáveis de entrada e saída, mas também o comportamento físico esperado para rochas anisotrópicas, reforçando sua adequação para a interpolação da resistência de pico dos filitos analisados.

No Apêndice 2.2 encontram-se detalhados os resultados complementares do modelo Random Forest, incluindo tabelas completas das métricas estatísticas por litotipo, gráficos de dispersão entre valores observados e preditos e representações polares das curvas $\sigma_1 - \sigma_3$ em função de β_{eff} .

4.3.2 Avaliação do modelo de *Gaussian Process Regression* (GPR)

A aplicação do modelo *Gaussian Process Regression* (GPR) teve como finalidade verificar se uma abordagem probabilística poderia melhorar a representação contínua da resistência de pico em função de β_{eff} e σ_3 . O desempenho foi avaliado com base nas mesmas métricas adotadas anteriormente, possibilitando comparação direta com o modelo *Random Forest*.

Os valores das métricas de desempenho apresentados na Tabela 8 indicam que o modelo GPR apresentou excelente capacidade de ajuste aos dados experimentais para a maioria dos litotipos analisados. Observa-se que os valores de R^2 no conjunto de teste são elevados para os filitos *Carbonaceous* (0,999), *Quartzite* (0,991), *Dry* (0,989) e *Moretown* (0,942), indicando adequada capacidade de generalização.

Tabela 8: Desempenho do modelo *Gaussian Process Regression* na predição da resistência de pico (σ_1).

Litotipo	R^2 (treino)	R^2 (teste)	Amplitude de σ_1 (MPa)	MAE (teste)	Erro médio	RMSE (teste)	Erro médio
<i>Carbonaceous</i>	1,000	0,999	272,8	1,804	1%	2,24	1%
<i>Quartzite</i>	1,000	0,991	279,4	5,249	2%	6,434	2%
<i>Micaceous</i>	0,983	0,702	327,3	25,947	8%	30,595	9%
<i>Dry</i>	0,997	0,989	294,1	5,09	2%	5,794	2%
<i>Moretown</i>	1,000	0,942	602,4	28,073	5%	37,399	6%

Fonte: (Autora, 2026).

Para os filitos *Carbonaceous*, *Quartzite* e *Dry*, os valores reduzidos de RMSE e MAE variam entre 1% e 2% da amplitude total de σ_1 , evidenciando elevada precisão do modelo. No caso do filito *Moretown*, embora a amplitude da resistência seja significativamente maior (≈ 602 MPa), os erros relativos permanecem moderados ($\approx 5\text{--}6\%$), representando melhoria significativa em relação ao modelo *Random Forest*.

Em contrapartida, o filito *Micaceous* apresentou desempenho inferior, com R^2 de 0,702 no conjunto de teste e erro relativo da ordem de 8–9%. Esse comportamento pode estar associado à maior variabilidade local e à presença de picos mais acentuados na curva experimental, que tendem a ser suavizados pelo modelo GPR, resultando em aumento do erro médio.

A Figura 10 apresenta a comparação entre valores experimentais e preditos. Observa-se forte alinhamento dos pontos à linha para a maioria dos litotipos, indicando boa concordância entre valores observados e estimados. O filito *Micaceous* apresenta maior dispersão no conjunto de teste, enquanto o filito *Moretown* demonstra melhoria significativa na aderência dos pontos em relação ao modelo anterior, evidenciando melhor captura do padrão global pelo GPR.

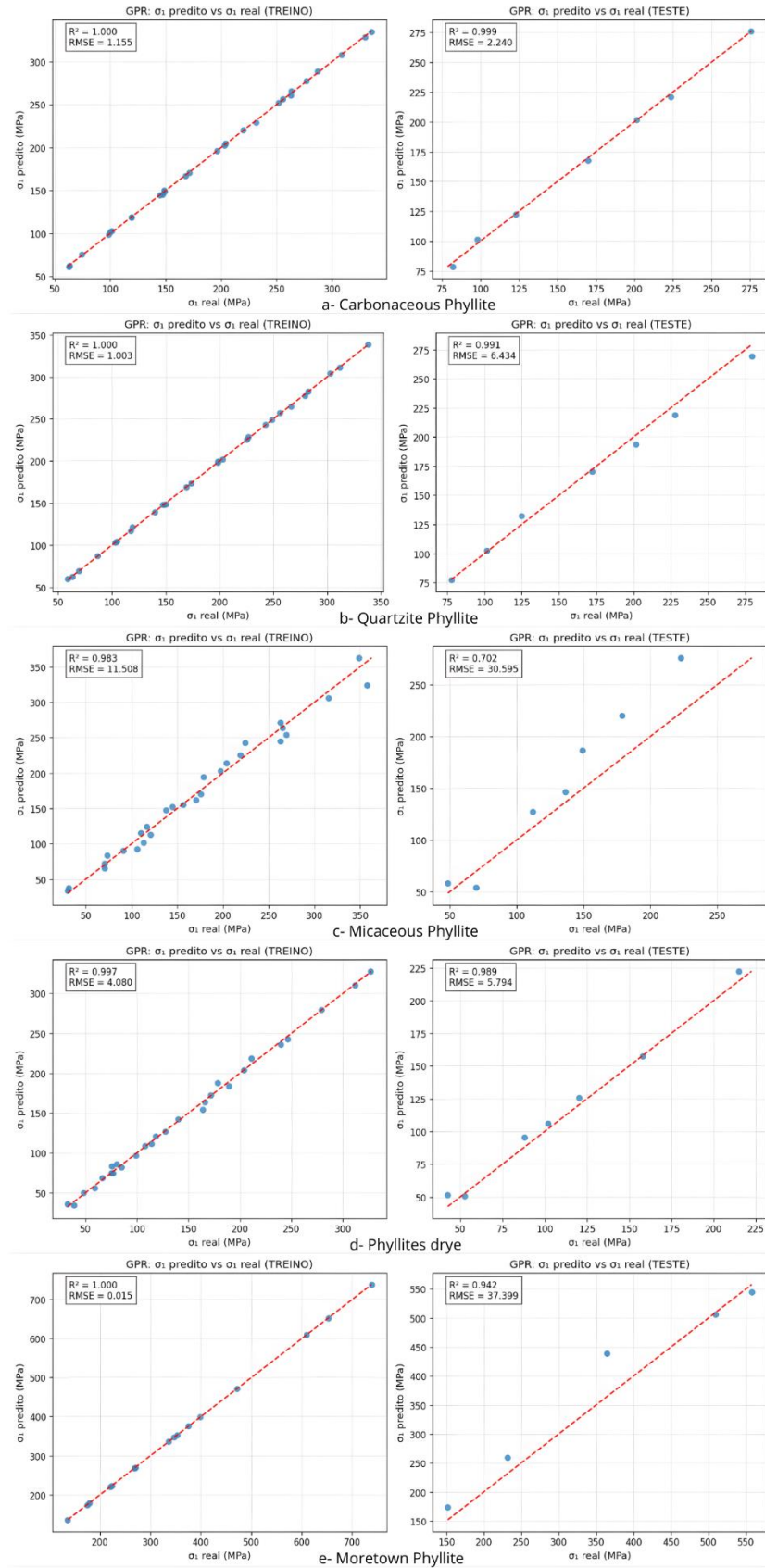


Figura 10: Comparação entre valores reais e preditos de σ_1 pelo modelo GPR.

Fonte: (Autora, 2026).

No Apêndice 2.3 são apresentados os resultados complementares do modelo *Gaussian Process Regression* (GPR), contemplando as métricas estatísticas detalhadas por litotipo, gráficos de valores reais versus preditos, curvas interpoladas e respectivas bandas de confiança associadas às estimativas do modelo.

As curvas interpoladas obtidas pelo modelo GPR apresentam comportamento suave e contínuo (Figura 11), reproduzindo o padrão anisotrópico em “U” identificado na análise exploratória. Observa-se preservação dos mínimos em ângulos intermediários e aumento global da resistência com o incremento da tensão confinante.

A suavidade característica do GPR favorece a representação contínua do comportamento mecânico, especialmente em regiões intermediárias do espaço de entrada.

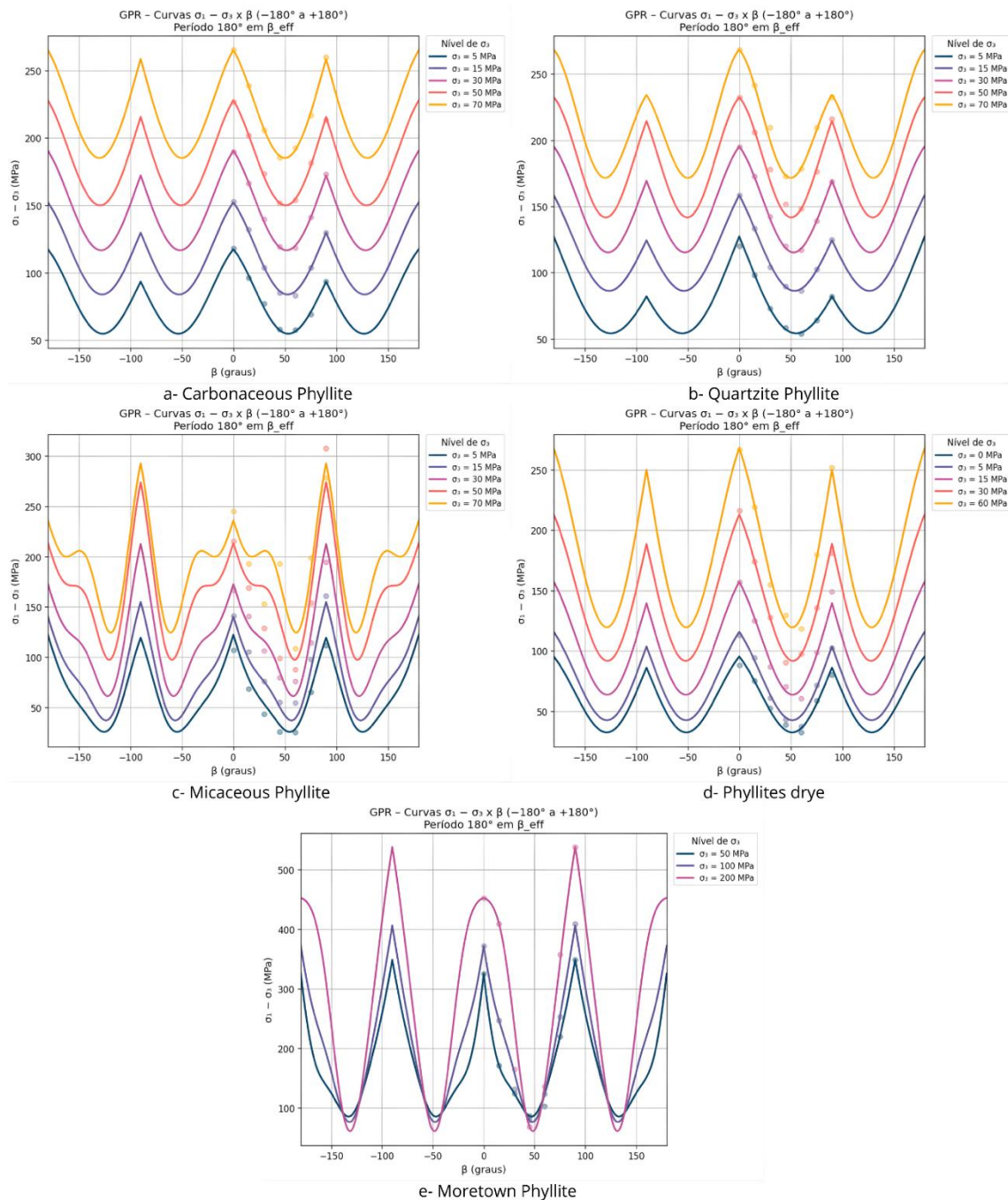


Figura 11: Curvas $\sigma_1 - \sigma_3 \times \beta_{eff}$ preditas pelo GPR

Fonte: (Autora, 2026).

As curvas de aprendizado apresentadas na Figura 12 indicam que, para os filitos *Carbonaceous*, *Quartzite* e *Dry*, há rápida convergência entre as curvas de treinamento e validação, com estabilização do R^2 próximo a 1,0 a partir de aproximadamente 10 a 15 amostras de treino, indicando boa capacidade de generalização e estabilidade do modelo.

Para o filito *Micaceous*, observa-se maior instabilidade nas fases iniciais, com diferença mais pronunciada entre as curvas de treinamento e validação, sugerindo maior sensibilidade à

quantidade de dados disponíveis e maior sensibilidade do modelo quando treinado com poucas amostras.

No caso do filito *Moretown*, embora o conjunto de dados seja reduzido, observa-se convergência progressiva das curvas à medida que o número de amostras aumenta, indicando que o modelo se beneficia da regularidade do padrão mecânico desse litotipo.

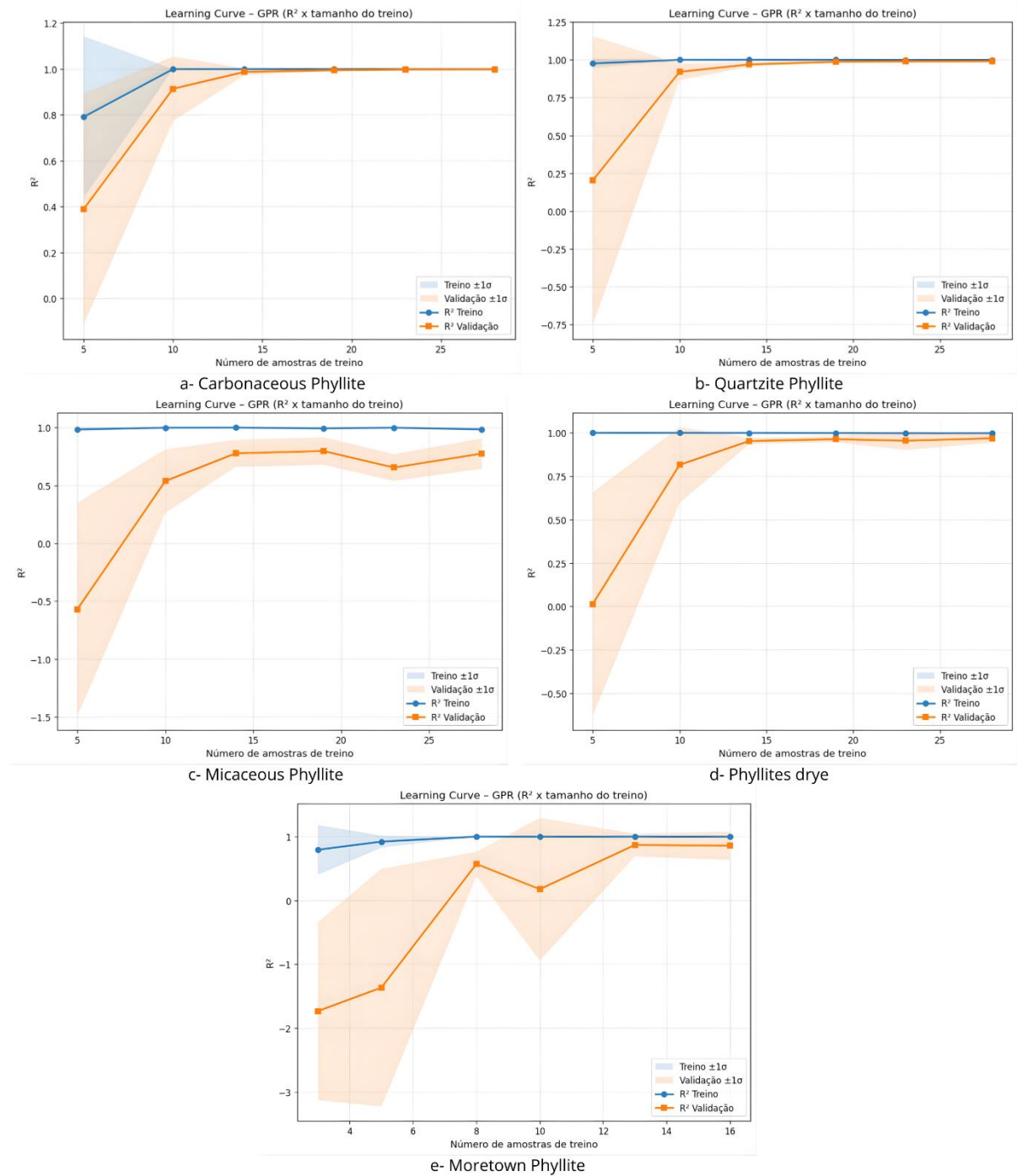


Figura 12: Curva de aprendizado (*Learning Curve*) para o modelo GPR.

Fonte: (Autora, 2026).

4.3.3 Avaliação do modelo de *Support Vector Regression* (SVR)

Por fim, o modelo *Support Vector Regression* (SVR) foi aplicado com o intuito de comparar seu desempenho com os modelos anteriormente analisados, mantendo os mesmos critérios de validação e métricas estatísticas. A análise concentrou-se na capacidade do modelo em capturar o padrão anisotrópico observado experimentalmente e na sua estabilidade frente às diferentes características dos bancos de dados.

Os resultados apresentados na Tabela 9 indicam desempenho satisfatório para a maioria dos litotipos analisados. Para os filitos *Carbonaceous*, *Quartzite*, *Micaceous* e *Dry*, os valores de R^2 no conjunto de teste foram elevados (0,998; 0,993; 0,967 e 0,970, respectivamente), enquanto os erros relativos permaneceram baixos, variando entre 1% e 3% da amplitude total de σ_1 . Esses resultados indicam boa capacidade de generalização e adequada representação da tendência global da resistência de pico.

Tabela 9: Desempenho do modelo SVR na predição da resistência de pico (σ_1).

Litotipo	R^2 (treino)	R^2 (teste)	Amplitude de σ_1 (MPa)	MAE (teste)	Erro médio	RMSE (teste)	Erro médio
<i>Carbonaceous</i>	0,998	0,998	272,8	2,765	1%	3,025	1%
<i>Quartzite</i>	1,000	0,993	279,4	4,601	2%	5,692	2%
<i>Micaceous</i>	0,952	0,967	327,3	8,273	3%	10,221	3%
<i>Dry</i>	0,994	0,970	294,1	7,380	3%	9,677	3%
<i>Moretown</i>	0,014	-0,005	602,4	139,791	23%	156,077	26%

Fonte: (Autora, 2026).

Em contrapartida, o filito *Moretown* apresentou desempenho insatisfatório, com R^2 de teste igual a $-0,005$. Um valor negativo de R^2 indica que o modelo apresentou desempenho insatisfatório, evidenciando incapacidade de capturar o comportamento mecânico desse litotipo. Além disso, os elevados valores de RMSE (≈ 156 MPa) e MAE (≈ 140 MPa), correspondendo a erros relativos entre 23% e 26% da amplitude total, confirmam a limitação do modelo nesse caso específico. Esse comportamento pode estar associado ao número reduzido de amostras disponíveis e à elevada amplitude da resistência observada experimentalmente.

Para verificar a possibilidade de sobreajuste, foi realizada validação cruzada do tipo K-fold (5 partições) para o modelo SVR, com embaralhamento dos dados e controle de reprodutibilidade. A análise do coeficiente de determinação médio entre os folds, bem como de seu desvio padrão, indicou estabilidade no desempenho do modelo. Esses resultados sugerem que os elevados valores de R^2 observados para determinados litotipos não decorrem

exclusivamente de memorização do conjunto de treinamento, mas refletem a capacidade do modelo em capturar padrões estruturais consistentes presentes na base de dados.

A análise dos gráficos de valores reais versus preditos (Figura 13) reforça essas conclusões. Para *Carbonaceous*, *Quartzite*, *Micaceous* e *Dry*, observa-se forte alinhamento dos pontos em torno da linha 1:1, indicando boa concordância entre valores observados e estimados. Para o filito *Moretown*, entretanto, observa-se dispersão significativa e ausência de aderência à linha, evidenciando baixa capacidade preditiva do modelo.

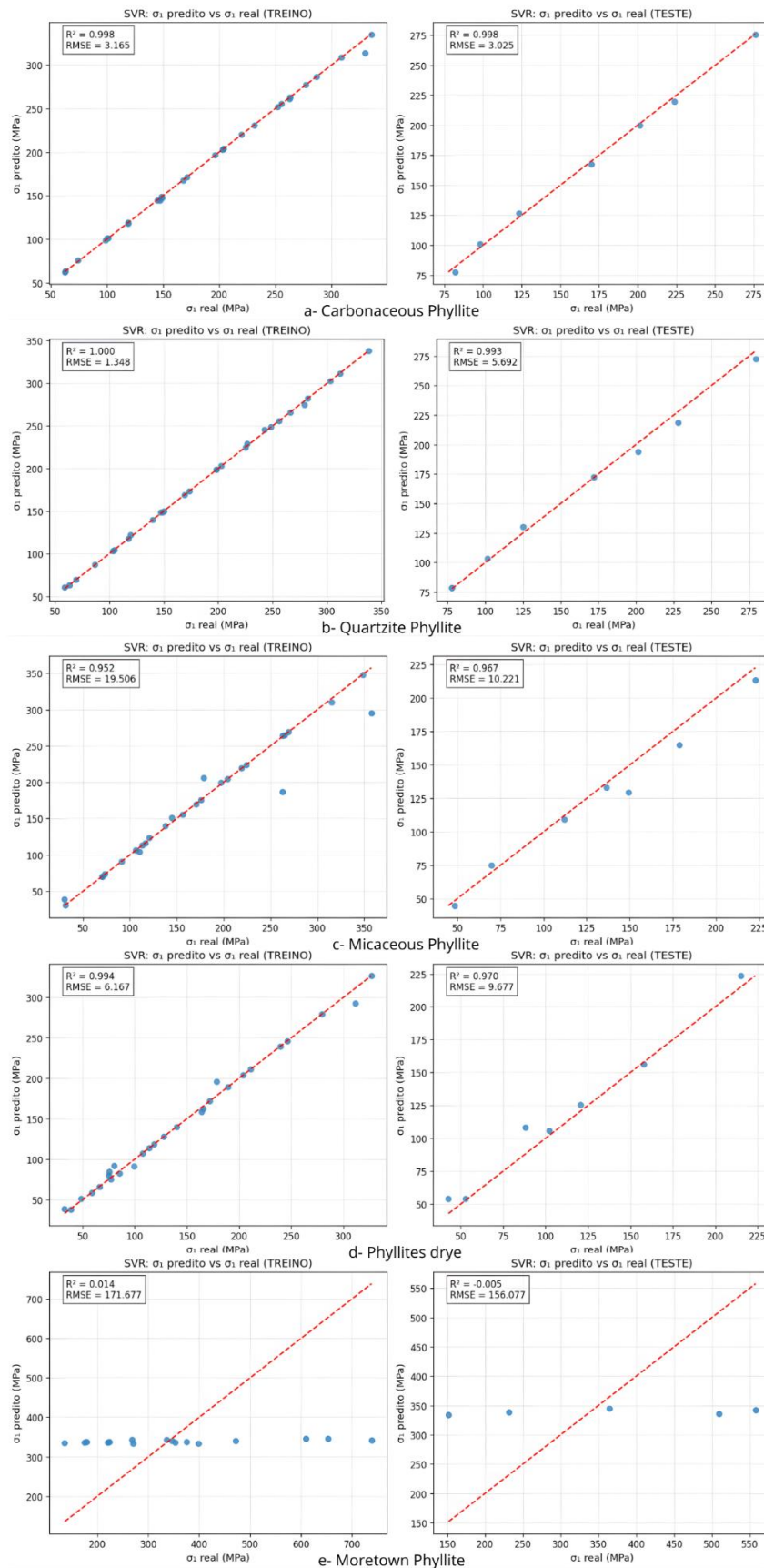


Figura 13: Comparação entre valores reais e preditos de σ_1 pelo modelo SVR.

Fonte: (Autora, 2026).

No Apêndice 2.4 encontram-se sumarizados os resultados adicionais do modelo *Support Vector Regression* (SVR), incluindo tabelas completas das métricas de desempenho, gráficos comparativos entre valores observados e estimados e representações das curvas preditas da resistência em função da orientação estrutural e da tensão confinante.

As curvas interpoladas $\sigma_1 - \sigma_3 \times \beta_{\text{eff}}$ obtidas pelo SVR (Figura 14) reproduzem o padrão anisotrópico em formato de “U” para a maioria dos litotipos, preservando os mínimos em ângulos intermediários e o aumento global da resistência com o incremento da tensão confinante. No caso do filito *Micaceous*, observa-se que o SVR apresenta representação mais robusta da tendência global quando comparado ao GPR, possivelmente devido à menor sensibilidade a pontos isolados.

Para o filito *Moretown*, entretanto, o modelo não reproduz adequadamente o padrão anisotrópico esperado, não apresentando claramente o formato característico em “U”. Essa ausência de coerência física indica limitação do SVR na representação do comportamento anisotrópico desse banco de dados específico.

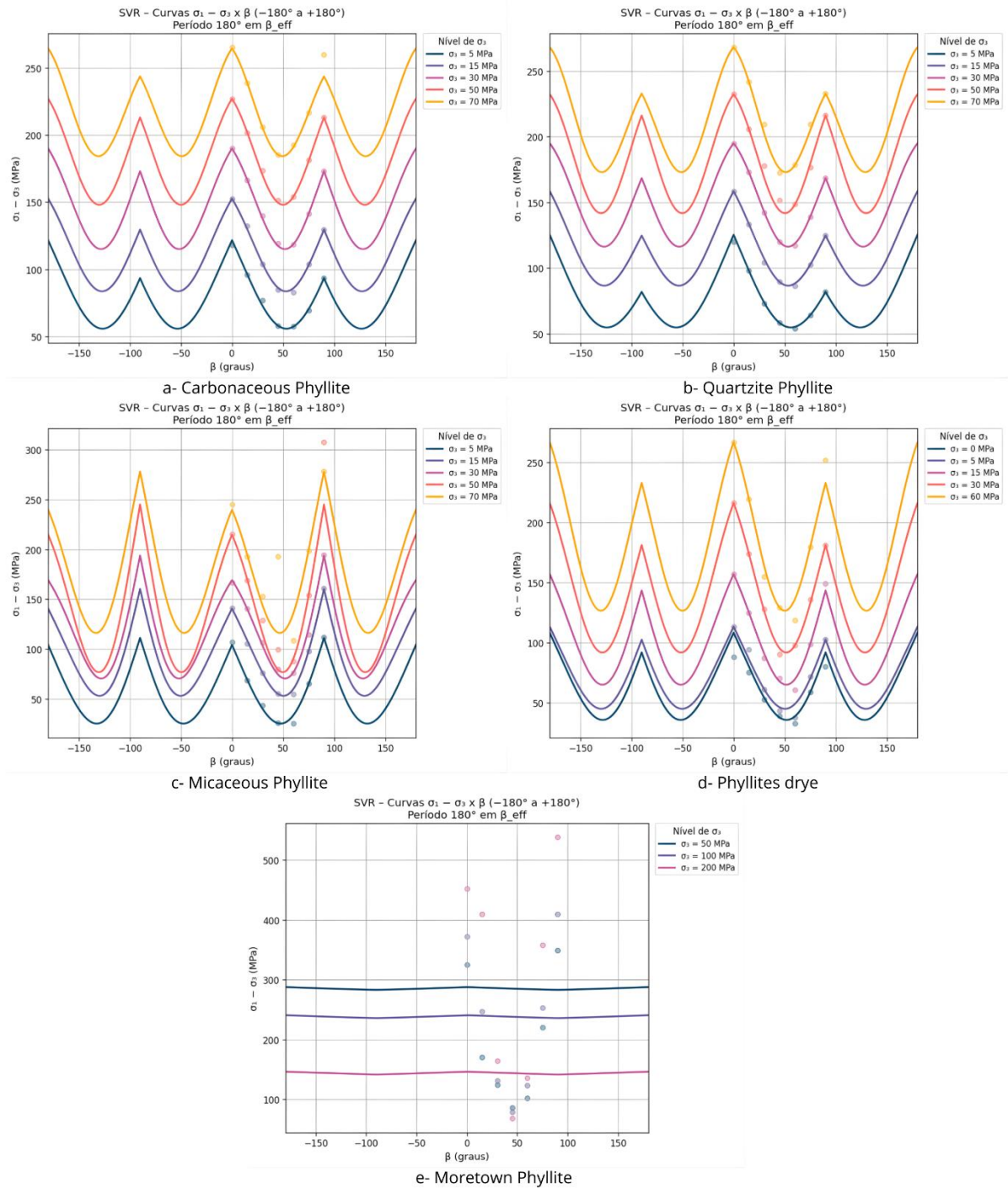


Figura 14: Curvas $\sigma_1 - \sigma_3 \times \beta_{eff}$ preditas pelo SVR.

Fonte: (Autora, 2026).

As curvas de aprendizado (Figura 15) indicam convergência progressiva entre as curvas de treinamento e validação para *Carbonaceous*, *Quartzite* e *Micaceous*, com estabilização do R^2 próximo a 1,0 a partir de aproximadamente 12 a 17 amostras de treino, indicando estabilidade e boa capacidade de generalização. Para o filito *Dry*, observa-se maior sensibilidade ao tamanho do conjunto de dados, com estabilização apenas para conjuntos

maiores. No caso do filito *Moretown*, as curvas de treino e validação não convergem de forma consistente, evidenciando instabilidade do modelo e elevada variância.

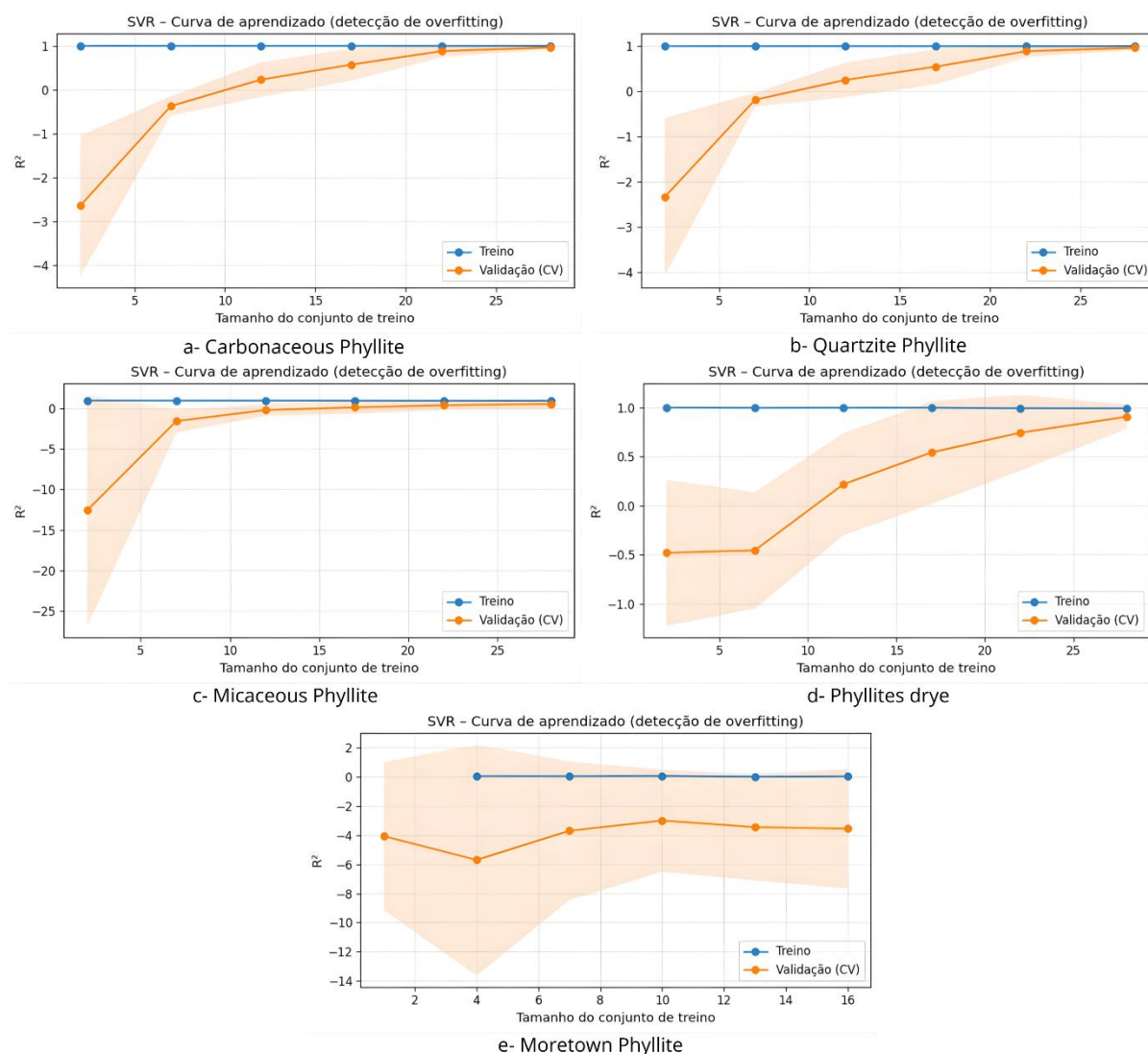


Figura 15: Curva de aprendizado (*Learning Curve*) para o modelo SVR.

Fonte: (Autora, 2026).

4.4 Comparação entre os modelos preditivos

A comparação quantitativa entre os modelos *Random Forest* (RF), *Gaussian Process Regression* (GPR) e *Support Vector Regression* (SVR) evidencia diferenças significativas no desempenho conforme o litotipo analisado (Tabela 10). De modo geral, o modelo GPR apresentou os maiores valores de R^2 no conjunto de teste e os menores erros relativos para a maioria dos filitos, destacando-se como o modelo mais preciso sob o ponto de vista estatístico. Para o filito *Micaceous*, entretanto, o modelo SVR apresentou melhor desempenho relativo,

evidenciado por maior R^2 no conjunto de teste e menor erro percentual, indicando melhor representação da tendência global desse banco de dados específico.

Tabela 10: Síntese comparativa dos modelos.

Litotipo	Modelo com maior R^2 (teste)	Modelo com menor erro relativo (%)	Observação
Carbonaceous	GPR	GPR	Excelente desempenho geral
Quartzite	GPR	GPR	Alta precisão
Micaceous	SVR	SVR	Melhor representação global
Dry	GPR	GPR	Alta estabilidade
Moretown	GPR	GPR	Melhor captura do padrão “U”

Fonte: (Autora, 2026).

O modelo Random Forest apresentou desempenho consistente para litotipos com menor variabilidade mecânica relativa, porém demonstrou limitação na generalização para o filito *Moretown*. O modelo SVR apresentou desempenho competitivo para os filitos *Carbonaceous*, *Quartzite* e *Dry*, mas evidenciou instabilidade significativa para o filito *Moretown*, refletida por valor negativo de R^2 no conjunto de teste e elevados erros relativos.

Sob o ponto de vista da coerência física, todos os modelos foram capazes de reproduzir o padrão anisotrópico característico dos filitos para a maioria dos litotipos, preservando o formato em “U” das curvas $\sigma_1 - \sigma_3 \times \beta_{\text{eff}}$ e representando adequadamente o aumento global da resistência com o incremento da tensão confinante. Contudo, para o filito *Moretown*, apenas o modelo GPR foi capaz de representar de forma consistente o comportamento anisotrópico esperado. O modelo SVR não reproduziu adequadamente o padrão em “U” para esse litotipo, evidenciando limitação na captura do comportamento mecânico sob elevada amplitude de resistência e número reduzido de amostras.

Para o filito *Micaceous*, por outro lado, o SVR apresentou melhor aderência à tendência global, mostrando-se menos sensível a oscilações locais presentes no conjunto experimental quando comparado ao GPR.

A análise das curvas de aprendizado indica que o modelo GPR apresentou convergência mais estável entre treino e validação para a maioria dos litotipos. O modelo SVR apresentou comportamento estável para conjuntos de dados intermediários, porém evidenciou maior sensibilidade à quantidade de amostras no caso do filito *Moretown*.

De forma geral, a comparação entre os modelos indica que o GPR apresentou o melhor desempenho global, tanto sob o ponto de vista estatístico quanto físico, destacando-se pela elevada precisão e pela capacidade de representar de forma contínua o comportamento

anisotrópico dos filitos analisados. O modelo SVR mostrou-se competitivo para a maioria dos litotipos, sendo particularmente eficaz na representação da tendência global em casos com maior irregularidade local. O modelo *Random Forest* apresentou desempenho consistente para litotipos com menor variabilidade mecânica relativa, porém com limitações na generalização para bancos de dados com ampla faixa de resistência e menor número de amostras.

Esses resultados reforçam a importância da escolha adequada do modelo preditivo em função das características mecânicas e estruturais do material analisado.

4.5 Síntese dos principais resultados

A análise desenvolvida neste capítulo permitiu avaliar de forma completa o comportamento anisotrópico dos filitos estudados e o desempenho dos modelos de aprendizado de máquina na interpolação da resistência de pico (σ_1).

Do ponto de vista mecânico, todos os litotipos analisados apresentaram comportamento anisotrópico característico, evidenciado pela variação sistemática da resistência de pico em função do ângulo de orientação anisotrópico (β_{eff}). As curvas $\sigma_1 - \sigma_3 \times \beta_{eff}$ assumiram formato típico em “U”, com valores mínimos em ângulos intermediários e máximos próximos às orientações paralela e perpendicular à direção da tensão principal maior. Observou-se ainda que o aumento da tensão confinante (σ_3) promove elevação global da resistência e redução relativa do contraste anisotrópico, embora sem eliminar completamente o controle estrutural.

A comparação entre os litotipos evidenciou diferenças significativas quanto à magnitude da resistência e ao grau de variabilidade mecânica. O filito *Moretown* apresentou a maior amplitude de resistência de pico, configurando o cenário mais desafiador para modelagem preditiva. O filito *Micaceous* demonstrou maior variabilidade local, enquanto *Carbonaceous* e *Quartzite* apresentaram comportamento mais regular e menor dispersão experimental relativa.

No que se refere ao desempenho dos modelos preditivos, o *Gaussian Process Regression* (GPR) apresentou, de modo geral, o melhor desempenho global, combinando elevados valores de R^2 no conjunto de teste, baixos erros relativos e elevada coerência física na reprodução do padrão anisotrópico. O *Support Vector Regression* (SVR) mostrou-se competitivo para a maioria dos litotipos, destacando-se na representação da tendência global do filito *Micaceous*, embora tenha apresentado instabilidade significativa para o filito *Moretown*. O *Random Forest* apresentou desempenho consistente para litotipos com menor

variabilidade relativa, porém com limitações na generalização em cenários de maior amplitude de resistência e menor número de amostras.

De forma geral, os resultados demonstram que os modelos avaliados são capazes de capturar a dependência conjunta entre orientação estrutural e tensão confinante na resistência de pico dos filitos. Entretanto, a escolha do modelo mais adequado depende das características específicas do banco de dados, especialmente da amplitude de resistência, do grau de anisotropia e da densidade amostral disponível.

Essa síntese consolida a base técnica para as conclusões gerais do trabalho, evidenciando que abordagens baseadas em aprendizado de máquina podem representar de forma consistente o comportamento mecânico anisotrópico de rochas foliadas, desde que consideradas as particularidades estruturais de cada litotipo.

5 CONSIDERAÇÕES FINAIS

Este trabalho teve como objetivo desenvolver e avaliar modelos preditivos capazes de interpolar a resistência de pico (σ_1) de filitos anisotrópicos, considerando simultaneamente os efeitos da orientação da foliação (β_{eff}) e da tensão confinante (σ_3). Critérios clássicos de resistência para rochas anisotrópicas frequentemente exigem elevado número de ensaios laboratoriais, apresentam resultados pontuais associados a orientações específicas e demandam interpretação mecânica complexa. Nesse contexto, a aplicação de modelos preditivos baseados em aprendizado de máquina foi proposta como alternativa para representar de forma contínua e integrada o comportamento mecânico dos filitos.

A análise exploratória confirmou, para todos os bancos de dados estudados, a existência de anisotropia direcional significativa. A resistência de pico apresentou dependência sistemática da orientação do plano de foliação e do nível de tensão confinante, assumindo curvas características em formato típico de “U”, com valores máximos de σ_1 próximos às orientações de 0° e 90° , e mínimos em ângulos intermediários. Observou-se ainda que o aumento da tensão confinante promove elevação global da resistência e redução relativa do contraste anisotrópico, embora sem eliminar completamente o controle estrutural.

A comparação entre os litotipos evidenciou diferenças relevantes quanto à magnitude da resistência, amplitude de variação, grau de anisotropia e densidade amostral. O filito *Moretown* destacou-se por apresentar maior amplitude de resistência e menor número de ensaios disponíveis, configurando o cenário mais desafiador para modelagem preditiva. O filito *Micaceous* apresentou maior variabilidade local, enquanto *Carbonaceous* e *Quartzite* demonstraram comportamento mais regular e menor dispersão experimental relativa. Tais diferenças refletem não apenas a variabilidade mecânica intrínseca, mas também características estruturais e condições experimentais específicas de cada banco de dados.

No que se refere ao desempenho dos modelos preditivos, verificou-se que todos foram capazes de representar o comportamento anisotrópico de forma satisfatória, embora com desempenhos distintos conforme o litotipo analisado. O modelo *Gaussian Process Regression* (GPR) apresentou o melhor desempenho geral, combinando elevados valores de R^2 no conjunto de teste, baixos erros relativos e maior coerência física na reprodução do padrão anisotrópico, destacando-se especialmente na modelagem do filito *Moretown*, mesmo sob menor densidade amostral e elevada amplitude de resistência.

O modelo *Support Vector Regression* (SVR) mostrou-se competitivo para a maioria dos litotipos, apresentando elevado desempenho estatístico e destacando-se na representação da tendência global do filito *Micaceous*. O modelo *Random Forest* apresentou desempenho consistente para litotipos com menor variabilidade relativa, porém com menor robustez em cenários de maior amplitude de resistência e menor quantidade de dados.

Os resultados obtidos demonstram que modelos de aprendizado de máquina supervisionado podem representar de forma contínua e coerente o comportamento mecânico anisotrópico de rochas foliadas, reduzindo a necessidade de extrapolações pontuais baseadas exclusivamente em ensaios laboratoriais discretos. A integração entre análise mecânica e modelagem preditiva mostrou-se eficaz na captura da dependência conjunta entre orientação estrutural e confinamento.

Como limitações do estudo, destaca-se o tamanho reduzido e a natureza do banco de dados utilizado, composto por resultados experimentais compilados da literatura, com número relativamente limitado de ensaios para alguns litotipos. A sensibilidade dos modelos à densidade e à distribuição amostral evidencia que a variabilidade estatística pode influenciar o desempenho preditivo, especialmente em cenários com menor representatividade experimental.

Essa condição pode impactar a capacidade de generalização dos modelos e restringir a extrapolação das conclusões para contextos geológicos distintos daqueles contemplados na base analisada. Além disso, os modelos foram treinados individualmente por litotipo, não sendo avaliada a construção de um modelo unificado capaz de capturar padrões físicos comuns entre diferentes materiais anisotrópicos.

5.1 Sugestões para Trabalhos Futuros

Como perspectivas de continuidade, recomenda-se a ampliação dos bancos de dados experimentais, a inclusão de parâmetros microestruturais e mineralógicos, a aplicação dos modelos a outros tipos de rochas anisotrópicas e a integração das predições a análises de estabilidade geotécnica em condições reais de campo.

Adicionalmente, recomenda-se investigar de forma mais aprofundada a natureza do aprendizado realizado pelos modelos desenvolvidos, buscando responder se os algoritmos estão capturando padrões específicos de cada litotipo ou se são capazes de apreender um comportamento físico mais geral associado ao fenômeno da anisotropia em rochas foliadas.

Estudos futuros podem explorar estratégias de treinamento conjunto entre diferentes litotipos e validações cruzadas interlitológicas, avaliando a capacidade de generalização física dos modelos e sua aplicabilidade em contextos distintos daqueles utilizados na calibração original.

De forma geral, conclui-se que a abordagem proposta representa um avanço na modelagem do comportamento mecânico anisotrópico de filitos, evidenciando o potencial do aprendizado de máquina como ferramenta complementar aos métodos clássicos de caracterização geomecânica.

REFERÊNCIAS BIBLIOGRÁFICAS

- BREIMAN, L. Random Forests. **Machine Learning**, v. 45, p. 5–32, out. 2001.
- CABRAL, B. R. **Processos Gaussianos para Aprendizado Supervisionado**. Florianópolis, Santa Catarina: Universidade Federal de Santa Catarina, 3 jun. 2021.
- CERRATO, S. A. R. **Modelagem da resistência de rochas anisotrópicas utilizando computação evolucionária**. Ouro Preto - MG: Universidade Federal de Ouro Preto, 16 ago. 2024.
- DONATH, F. A. Effects of Cohesion and Granularity on Deformational Behavior of Anisotropic Rock. *Em: Geological Society of America*. Boulder, Colorado: Geological Society of America, 1972. p. 95–129.
- ESCOVEDO, TATIANA.; KOSHIYAMA, ADRIANO. **Introdução a data science : algoritmos de machine learning e métodos de análise**. [s.l.] Casa do Código, 2020.
- GOMES, G. J. C. *et al.* Bayesian inference of rock strength anisotropy: Uncertainty analysis of the Hoek–Brown failure criterion. **International Journal of Rock Mechanics and Mining Sciences**, v. 148, p. 104952, 1 dez. 2021.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. [s.l.: s.n.].
- JAEGER, J. C. Shear Failure of Anisotropic Rocks. **Geological Magazine**, v. 97, n. 1, p. 65–72, 1960.
- KUMAR, A. **Engineering Behaviour of Anisotropic Rocks**. [s.l.] Indian Institute of Technology Roorkee, jul. 2006.
- LAREDO SICSÚ, A.; LERNER BARTH, N. **TÉCNICAS DE MACHINE LEARNING**. [s.l.: s.n.].
- LIAW, A.; WIENER, M. Classification and Regression by randomForest. **R News**, v. 2/3, n. 3, p. 18–22, out. 2002.
- NASSERI, M. H. B.; RAO, K. S.; RAMAMURTHY, T. Anisotropic strength and deformational behavior of Himalayan schists. **International Journal of Rock Mechanics and Mining Sciences**, v. 40, n. 1, p. 3–23, jan. 2003.

PARK, J. *et al.* Gaussian Process Regression (GPR) Representation in Predictive Model Markup Language (PMML). **Smart and Sustainable Manufacturing Systems**, v. 1, n. 1, p. 121–142, 29 mar. 2017.

RAMAMURTHY, T.; VENKATAPPA RAO, G.; SINGH J. **A Strength Criterion for Anisotropic Rocks** Sydney: Fifth Australia–New Zealand Conference on Geomechanics, 23 ago. 1988

RAMAMURTHY, T.; VENKATAPPA RAO, G.; SINGH, J. Engineering behaviour of phyllites. **Engineering Geology**, v. 33, p. 209–225, 1993.

RAMAMURTHY, T.; VENKATAPPA RAO, G.; SINGH, J. Engineering behaviour of phyllites. **Engineering Geology**, v. 33, p. 209–225, 1993.

SCIKIT-LEARN. **Scikit-learn**. Disponível em: <<https://scikit-learn.org/stable/index.html>>. Acesso em: 10 fev. 2026.

SINGH, M. *et al.* A nonlinear criterion for triaxial strength of inherently anisotropic rocks. **Rock Mechanics and Rock Engineering**, v. 48, n. 4, p. 1387–1405, 21 jul. 2015.

SINGH, V. K.; SINGH, D.; SINGH, T. N. Prediction of strength properties of some schistose rocks from petrographic properties using artificial neural networks. **International Journal of Rock Mechanics & Mining Sciences**, v. 38, p. 269–284, 2001.

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, v. 14, n. 3, p. 199–222, ago. 2004.

SMOLA, A.; SMOLA, A. **Support Vector Machines and Kernel Algorithms** 2002 Disponível em: <<https://api.semanticscholar.org/CorpusID:13989473>>

VIKRAMADITYA JAKKULA. **Tutorial on Support Vector Machine (SVM)**. [s.l.] Arizona State University, 2006.

WANG, J. An Intuitive Tutorial to Gaussian Process Regression. 28 jan. 2024.

APÊNDICES

1. Código em *Python*:

Parte A – IMPORTS, CONFIGURAÇÃO GLOBAL E FUNÇÕES

```
# -*- coding: utf-8 -*-
"""Untitled3.ipynb

Automatically generated by Colab.

Original file is located at
    https://colab.research.google.com/drive/1sSc045XbEXNbuJ1I6zp5pgbId0
Hfp3QY
"""

# -*- coding: utf-8 -*-
"""
SCRIPT UNIFICADO - ANISOTROPIA (Partes A-H + E2)
Modelos: Random Forest, GPR, SVR (SVM Regressor com Tuning)
Curvas  $\sigma_1$ - $\sigma_3$  x  $\beta$ , gráficos polares, equações pares e estatística
descritiva.
Usa  $\beta_{eff}$  (dobramento 180°) como variável angular principal.
"""

# =====
#  Parte A - IMPORTS, CONFIGURAÇÃO GLOBAL E FUNÇÕES
#  =====

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

from pathlib import Path

from sklearn.model_selection import (
    train_test_split,
    cross_val_score,
    GridSearchCV,
    KFold,
    learning_curve,
)
from sklearn.ensemble import RandomForestRegressor
from sklearn.gaussian_process import GaussianProcessRegressor
```

```

from sklearn.gaussian_process.kernels import RBF, WhiteKernel,
ConstantKernel as C
from sklearn.svm import SVR
from sklearn.metrics import r2_score, mean_squared_error,
mean_absolute_error
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import Pipeline

from scipy.stats import norm
from IPython.display import display

# Configurações gerais de plot
plt.rcParams["figure.dpi"] = 120
plt.rcParams["axes.titlesize"] = 12
plt.rcParams["axes.labelsize"] = 11

# Diretório para salvar figuras
BASE_FIG_DIR = Path("figuras")
BASE_FIG_DIR.mkdir(exist_ok=True)

FIGS = [] # lista global de figuras (se quiser usar depois)

def salvar_fig_atual(nome_arquivo: str):
    """
    Salva a figura atual em PNG (300 dpi) na pasta 'figuras'
    e registra na lista global FIGS.
    """
    fig = plt.gcf()
    caminho = BASE_FIG_DIR / f"{nome_arquivo}.png"
    fig.savefig(caminho, dpi=300, bbox_inches="tight")
    FIGS.append((fig, nome_arquivo))
    plt.show()

def fold_180_deg(beta_deg):
    """
    Dobra o ângulo  $\beta$  em período 180°.
    Resultado:  $\beta_{\text{eff}} \in [0^\circ, 90^\circ]$ .
    """
    beta_eff = np.abs(((beta_deg + 90) % 180) - 90)
    return beta_eff

def style_tabela_cinza(df_est, caption):
    """
    Estilização padrão (tabela cinza) usada em todas as tabelas
    principais.
    """

```

```

return (
    df_est.round(3)
    .style
    .set_caption(caption)
    .set_table_styles([
        {
            "selector": "caption",
            "props": [
                ("caption-side", "top"),
                ("text-align", "center"),
                ("font-size", "15px"),
                ("font-weight", "bold"),
                ("margin-bottom", "10px"),
                ("color", "black"),
            ],
        },
        {
            "selector": "th",
            "props": [
                ("background-color", "#ececec"),
                ("text-align", "center"),
                ("font-size", "12px"),
                ("font-weight", "bold"),
                ("padding", "6px"),
            ],
        },
        {
            "selector": "td",
            "props": [
                ("text-align", "center"),
                ("font-size", "12px"),
                ("padding", "5px"),
            ],
        },
    ])
)

```

```

def avaliar_modelo(nome, modelo, X_train_s, X_test_s, y_train, y_test,
nome_fig_base):

```

```

    """
    Ajusta métricas de desempenho e gera gráfico de dispersão
    y_pred vs y_real para treino e teste.
    Retorna dicionário com R², RMSE e MAE para treino e teste.
    """

    # Previsões
    y_pred_train = modelo.predict(X_train_s)
    y_pred_test  = modelo.predict(X_test_s)

```

```

# Métricas
r2_train = r2_score(y_train, y_pred_train)
r2_test  = r2_score(y_test,  y_pred_test)
rmse_train = np.sqrt(mean_squared_error(y_train, y_pred_train))
rmse_test  = np.sqrt(mean_squared_error(y_test,  y_pred_test))
mae_train  = mean_absolute_error(y_train, y_pred_train)
mae_test   = mean_absolute_error(y_test,  y_pred_test)

print(f"\n==== {nome} =====")
print(f"Treino: R² = {r2_train:.3f} | RMSE = {rmse_train:.3f} | MAE = {mae_train:.3f}")
print(f"Teste : R² = {r2_test:.3f} | RMSE = {rmse_test:.3f} | MAE = {mae_test:.3f}")

# Gráfico de dispersão
fig, axes = plt.subplots(1, 2, figsize=(12, 5))

# --- Treino ---
ax = axes[0]
ax.scatter(y_train, y_pred_train, alpha=0.8)
min_val = min(y_train.min(), y_pred_train.min())
max_val = max(y_train.max(), y_pred_train.max())
ax.plot([min_val, max_val], [min_val, max_val], "r--")
ax.set_title(f"{nome}:  $\sigma_1$  predito vs  $\sigma_1$  real (TREINO)")
ax.set_xlabel(" $\sigma_1$  real (MPa)")
ax.set_ylabel(" $\sigma_1$  predito (MPa)")
ax.grid(alpha=0.3)
ax.text(
    0.05, 0.90,
    f"R² = {r2_train:.3f}\nRMSE = {rmse_train:.3f}",
    transform=ax.transAxes,
    fontsize=10,
    bbox=dict(facecolor="white", alpha=0.7),
)

# --- Teste ---
ax = axes[1]
ax.scatter(y_test, y_pred_test, alpha=0.8)
min_val = min(y_test.min(), y_pred_test.min())
max_val = max(y_test.max(), y_pred_test.max())
ax.plot([min_val, max_val], [min_val, max_val], "r--")
ax.set_title(f"{nome}:  $\sigma_1$  predito vs  $\sigma_1$  real (TESTE)")
ax.set_xlabel(" $\sigma_1$  real (MPa)")
ax.set_ylabel(" $\sigma_1$  predito (MPa)")
ax.grid(alpha=0.3)
ax.text(
    0.05, 0.90,
    f"R² = {r2_test:.3f}\nRMSE = {rmse_test:.3f}",

```

```

        transform=ax.transAxes,
        fontsize=10,
        bbox=dict(facecolor="white", alpha=0.7),
    )

plt.tight_layout()
salvar_fig_atual(f"{nome_fig_base}_scatter_sigma1_pred_vs_real")

return {
    "r2_train": r2_train,
    "r2_test": r2_test,
    "rmse_train": rmse_train,
    "rmse_test": rmse_test,
    "mae_train": mae_train,
    "mae_test": mae_test,
}

```

--- Parte B – CARREGAMENTO DO BANCO DE DADOS ---

```

from google.colab import files

print("📁 Selecione o arquivo Excel do banco de dados (.xlsx):")
uploaded = files.upload()
excel_name = list(uploaded.keys())[0]
print(f"✓ Arquivo recebido: {excel_name}")

df = pd.read_excel(excel_name)

print("\nPreview do DataFrame carregado:")
display(df.head())

# -----
# PRÉ-PROCESSAMENTO PARA ANISOTROPIA
# -----

# 1)  $|\beta|$ 
df["beta_abs"] = df["beta"].abs()

# 2)  $\beta_{\text{eff}}$  - variável física de anisotropia (período 180°)
df["beta_eff"] = fold_180_deg(df["beta"])

# 3) Combinações mecânicas
df["sigma_diff"] = df["sigma1"] - df["sigma3"]
df["sigma_ratio"] = np.where(df["sigma3"] != 0,
                             df["sigma1"] / df["sigma3"],

```

```

np.nan)

# 4) níveis de  $\sigma_3$ 
sigmas3_vals =
sorted(df["sigma3"].dropna().astype(int).unique().tolist())
print("\nNíveis de  $\sigma_3$  encontrados:", sigmas3_vals)

# -----
# FEATURES OFICIAIS PARA MODELAGEM
# -----

feature_cols = ["beta_eff", "r", "sigmac", "sigma3"]

X = df[feature_cols].values
y = df["sigma1"].values

print("\nFormato de X:", X.shape)
print("Formato de y:", y.shape)

# -----
# TREINO / TESTE COM ESTRATIFICAÇÃO EM  $\sigma_3$ 
# -----

X_train, X_test, y_train, y_test = train_test_split(
    X, y,
    test_size=0.20,
    random_state=42,
    stratify=df["sigma3"],
)

# Paleta fixa para os níveis de  $\sigma_3$ 
COLOR_SIGMA3 = ["#003f5c", "#58508d", "#bc5090", "#ff6361", "#ffa600"]

scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

print("\nTamanhos dos conjuntos:")
print("X_train:", X_train.shape)
print("X_test :", X_test.shape)

# Conjunto completo para CV
X_cv = df[feature_cols].values
y_cv = df["sigma1"].values

```

Parte C – ANÁLISE EXPLORATÓRIA DA ANISOTROPIA

```
# =====
# IMPORTS
# =====
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# =====
# VERIFICAÇÃO DO DATAFRAME
# =====
if "df" not in globals():
    raise RuntimeError(
        "O DataFrame 'df' não existe. "
        "Execute PRIMEIRO a célula onde você cria/carrega o df."
    )

# =====
# DEFINIÇÃO DOS NÍVEIS DE  $\sigma_3$ 
# =====
sigmas3_vals = np.sort(df["sigma3"].unique())

# =====
# SCATTER  $\sigma_1$  vs  $\beta_{eff}$  por nível de  $\sigma_3$ 
# =====
plt.figure(figsize=(8, 6))
for s3 in sigmas3_vals:
    sub = df[df["sigma3"] == s3]
    plt.scatter(
        sub["beta_eff"],
        sub["sigma1"],
        label=f" $\sigma_3 = \{s3\}$  MPa",
        alpha=0.7
    )

plt.xlabel(" $\beta_{eff}$  (graus, período 180°)")
plt.ylabel(" $\sigma_1$  (MPa)")
plt.title(" $\sigma_1$  vs  $\beta_{eff}$  por nível de  $\sigma_3$ ")
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()
salvar_fig_atual("C_sigma1_vs_betaeff_por_sigma3")
```

```

# =====
# SCATTER ( $\sigma_1 - \sigma_3$ ) vs  $\beta_{\text{eff}}$  por nível de  $\sigma_3$ 
# =====
plt.figure(figsize=(8, 6))
for s3 in sigmas3_vals:
    sub = df[df["sigma3"] == s3]
    plt.scatter(
        sub["beta_eff"],
        sub["sigma_diff"],
        label=f" $\sigma_3 = \{s3\}$  MPa",
        alpha=0.7
    )

plt.xlabel(" $\beta_{\text{eff}}$  (graus, período 180°)")
plt.ylabel(" $\sigma_1 - \sigma_3$  (MPa)")
plt.title(" $\sigma_1 - \sigma_3$  vs  $\beta_{\text{eff}}$  por nível de  $\sigma_3$ ")
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()
salvar_fig_atual("C_sigmadiff_vs_betaeff_por_sigma3")

# =====
# QUANTIFICAÇÃO DA ANISOTROPIA
#  $A = \sigma_{\text{max}} / \sigma_{\text{min}}$ 
# =====
anisotropy_rows = []

for s3 in sigmas3_vals:
    sub = df[df["sigma3"] == s3]

    # --- definição clássica
    sigma_max = sub["sigma1"].max()
    sigma_min = sub["sigma1"].min()
    A_raw = sigma_max / sigma_min if sigma_min != 0 else np.nan

    # --- definição robusta (mediana por orientação  $\beta_{\text{eff}}$ )
    by_beta = sub.groupby("beta_eff")["sigma1"].median()
    A_beta_med = (
        by_beta.max() / by_beta.min()
        if by_beta.min() != 0 else np.nan
    )

    anisotropy_rows.append({
        "sigma3": s3,
        "A_raw": A_raw,
        "A_beta_med": A_beta_med
    })

```

```

anisotropy_df = pd.DataFrame(anisotropy_rows)
display(anisotropy_df)

# =====
# GRÁFICO DO GRAU DE ANISOTROPIA
# =====
plt.figure(figsize=(8, 5))
plt.plot(
    anisotropy_df["sigma3"],
    anisotropy_df["A_beta_med"],
    marker="o"
)

plt.xlabel("σ3 (MPa)")
plt.ylabel("A = σmax / σmin")
plt.title("Grau de anisotropia por nível de σ3")
plt.grid(alpha=0.3)
plt.tight_layout()
salvar_fig_atual("C_anisotropia_A_por_sigma3")
=====

```

Parte D – MODELO 1: RANDOM FOREST

```

=====

rf = RandomForestRegressor(
    n_estimators=800,
    max_depth=None,
    min_samples_split=2,
    min_samples_leaf=1,
    max_features="sqrt",
    bootstrap=True,
    random_state=42,
)

rf.fit(X_train_scaled, y_train)

metrics_rf = avaliar_modelo(
    nome="Random Forest",
    modelo=rf,
    X_train_s=X_train_scaled,
    X_test_s=X_test_scaled,
    y_train=y_train,
    y_test=y_test,
    nome_fig_base="D_RF"
)

# Validação cruzada 5-fold

```

```

rf_cv_model = Pipeline([
    ("scaler", StandardScaler()),
    ("rf", RandomForestRegressor(
        n_estimators=800,
        max_depth=None,
        min_samples_split=2,
        min_samples_leaf=1,
        max_features="sqrt",
        bootstrap=True,
        random_state=42,
    ))
])

rf_cv_scores = cross_val_score(
    rf_cv_model,
    X_cv, y_cv,
    cv=5,
    scoring="r2"
)

print(f"\n[Random Forest] CV 5-fold -> R² médio =
{rf_cv_scores.mean():.3f} ± {rf_cv_scores.std():.3f}")

# =====
# 4) CURVAS  $\sigma_1 - \sigma_3 \times \beta$  ( $-180^\circ$  a  $+180^\circ$ ) - RF
# =====

def gerar_curvas_e_equacoes_pares_rf():
    """
    Curvas ( $\sigma_1 - \sigma_3$ ) x  $\beta$  em  $\beta \in [-180, 180]$  com período  $180^\circ$ 
    e tabela com as equações pares completas (Random Forest).
    """
    fig, ax = plt.subplots(figsize=(10, 6))
    equations = []

    for i, s3 in enumerate(sigmas3_vals):
        sub = df[df["sigma3"] == s3].copy()
        if sub.empty:
            continue

        cor = COLOR_SIGMA3[i % len(COLOR_SIGMA3)]

        # Pontos reais
        y_real = sub["sigma1"] - s3
        ax.scatter(sub["beta"], y_real, alpha=0.4, s=30, color=cor)

        # Curva base RF em  $\beta_{\text{eff}}$  de 0 a 90
        beta_eff_base = np.linspace(0, 90, 181)

```

```

r_ref      = sub["r"].mean()
sigmac_ref = sub["sigmac"].mean()

grid_base = pd.DataFrame({
    "beta_eff": beta_eff_base,
    "r":      r_ref,
    "sigmac":  sigmac_ref,
    "sigma3":  s3,
})

grid_base_scaled =
scaler.transform(grid_base[feature_cols].values)
sigma1_base = rf.predict(grid_base_scaled)
y_base = sigma1_base - s3

# Ajuste PAR COMPLETO:  $f(\beta_{\text{eff}}) = a_4 \beta_{\text{eff}}^4 + a_2 \beta_{\text{eff}}^2 + a_0$ 
z = beta_eff_base**2
A, B, Cc = np.polyfit(z, y_base, deg=2)
a4 = A
a2 = B
a0 = Cc

eq_str = (
    f"{a4:.4e} ·  $\beta_{\text{eff}}^4$  "
    f"+ {a2:.4e} ·  $\beta_{\text{eff}}^2$  "
    f"+ {a0:.4e}"
).replace("+ -", "- ")
equations.append((s3, a4, a2, a0, eq_str))

# Curva RF estendida em  $\beta \in [-180, 180]$ 
beta_full = np.linspace(-180, 180, 721)
beta_eff_full = fold_180_deg(beta_full)

grid_full = pd.DataFrame({
    "beta_eff": beta_eff_full,
    "r":      r_ref,
    "sigmac":  sigmac_ref,
    "sigma3":  s3,
})

grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
sigma1_full = rf.predict(grid_full_scaled)
y_full = sigma1_full - s3

# Só a curva entra na legenda
ax.plot(
    beta_full,

```

```

        y_full,
        linewidth=2,
        color=cor,
        label=f" $\sigma_3 = \{s3\}$  MPa"
    )

ax.set_xlabel(" $\beta$  (graus)")
ax.set_ylabel(" $\sigma_1 - \sigma_3$  (MPa)")
ax.set_title(
    "Random Forest - Curvas  $\sigma_1 - \sigma_3$  x  $\beta$  (-180° a +180°)\n"
    "Período 180° em  $\beta_{eff}$ "
)
ax.grid(True)
ax.set_xlim(-180, 180)

# Legenda ao lado
ax.legend(
    title="Nível de  $\sigma_3$ ",
    loc="upper left",
    bbox_to_anchor=(1.02, 1.0),
    borderaxespad=0.0,
    fontsize=9,
    frameon=True
)

fig.tight_layout(rect=[0, 0, 0.80, 1])
salvar_fig_atual("D_RF_curvas_sigma_diff_perodo_180")

# Tabela das equações
eq_df = pd.DataFrame(
    [(s3, a4, a2, a0, eq_str) for (s3, a4, a2, a0, eq_str) in
equations],
    columns=[" $\sigma_3$  (MPa)", "a4", "a2", "a0", "Equação ( $\sigma_1 -$   

 $\sigma_3$ ) ( $\beta_{eff}$ )"]
)
display(style_tabela_cinza(eq_df, "Equações pares - Random
Forest"))

def plot_polar_360_sigma_diff_rf():
    fig = plt.figure(figsize=(7, 7))
    ax = fig.add_subplot(111, projection="polar")

    fig.subplots_adjust(right=0.80)

    ax.set_theta_zero_location("N")
    ax.set_theta_direction(-1)
    ax.set_thetamin(0)
    ax.set_thetamax(360)

```

```

ax.set_thetagrids(range(0, 361, 30))

colors = plt.cm.viridis(np.linspace(0, 1, len(sigmas3_vals)))
handles = []

for i, s3 in enumerate(sigmas3_vals):
    sub = df[df["sigma3"] == s3].copy()
    if sub.empty:
        continue

    beta_real = sub["beta"].values
    theta_real = np.deg2rad((beta_real + 360) % 360)
    r_real = (sub["sigma1"] - s3).values

    h1 = ax.scatter(theta_real, r_real, s=40,
                    color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
alpha=0.85)
    handles.append((h1, f"Real  $\sigma_3$  = {s3} MPa"))

    beta_full = np.linspace(0, 360, 721)
    beta_eff_full = fold_180_deg(beta_full)

    r_ref = sub["r"].mean()
    sigmac_ref = sub["sigmac"].mean()

    grid_full = pd.DataFrame({
        "beta_eff": beta_eff_full,
        "r": r_ref,
        "sigmac": sigmac_ref,
        "sigma3": s3,
    })

    grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
    sigma1_pred = rf.predict(grid_full_scaled)
    r_pred = sigma1_pred - s3

    theta_full = np.deg2rad(beta_full)
    h2, = ax.plot(theta_full, r_pred,
                    color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
linestyle="--", linewidth=2)
    handles.append((h2, f"Pred  $\sigma_3$  = {s3} MPa (RF)"))

    ax.set_title(r"( $\sigma_1 - \sigma_3$ ) vs  $\beta$  - Random Forest (0°-360°, período
180°)",
                fontsize=13, pad=30)

ax.grid(alpha=0.3)

```

```

h_objs, labels = zip(*handles)
ax.legend(h_objs, labels,
          loc="upper left", bbox_to_anchor=(1.02, 1.0),
          borderaxespad=0.0, fontsize=9, frameon=True)

plt.tight_layout()
salvar_fig_atual("D_RF_polar_360_sigma_diff")

# =====
# 6) POLAR:  $\sigma_1/\sigma_3 \times \beta - 0^\circ$  a  $360^\circ$  - RF
# =====

def plot_polar_360_sigma_ratio_rf():
    fig = plt.figure(figsize=(7, 7))
    ax = fig.add_subplot(111, projection="polar")

    fig.subplots_adjust(right=0.80)

    ax.set_theta_zero_location("N")
    ax.set_theta_direction(-1)
    ax.set_thetamin(0)
    ax.set_thetamax(360)
    ax.set_theta grids(range(0, 361, 30))

    sigmas3_vals_ratio = [s for s in sigmas3_vals if s != 0]

    colors = plt.cm.plasma(np.linspace(0, 1, len(sigmas3_vals_ratio)))
    handles = []

    for i, s3 in enumerate(sigmas3_vals_ratio):
        sub = df[df["sigma3"] == s3].copy()
        if sub.empty:
            continue

        beta_real = sub["beta"].values
        theta_real = np.deg2rad((beta_real + 360) % 360)
        r_real = (sub["sigma1"] / s3).values

        h1 = ax.scatter(theta_real, r_real, s=40,
                        color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
                        alpha=0.85)
        handles.append((h1, f"Real  $\sigma_3 = \{s3\}$  MPa"))

        beta_full = np.linspace(0, 360, 721)
        beta_eff_full = fold_180_deg(beta_full)

        r_ref = sub["r"].mean()

```

```

sigma_ref = sub["sigma"].mean()

grid_full = pd.DataFrame({
    "beta_eff": beta_eff_full,
    "r":        r_ref,
    "sigma":    sigma_ref,
    "sigma3":   s3,
})

grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
sigma_pred = rf.predict(grid_full_scaled)
r_pred      = sigma_pred / s3

theta_full = np.deg2rad(beta_full)
h2, = ax.plot(theta_full, r_pred,
               color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
               linestyle="--", linewidth=2)
handles.append((h2, f"Pred  $\sigma_3$  = {s3} MPa (RF)"))

ax.set_title(r" $\sigma_1/\sigma_3$  vs  $\beta$  - Random Forest (0°-360°, período 180°)",
             fontsize=13, pad=30)

ax.grid(alpha=0.3)

h_objs, labels = zip(*handles)
ax.legend(h_objs, labels,
          loc="upper left", bbox_to_anchor=(1.02, 1.0),
          borderaxespad=0.0, fontsize=9, frameon=True)

plt.tight_layout()
salvar_fig_atual("D_RF_polar_360_sigma_ratio")

# Chamada das funções RF
gerar_curvas_e_equacoes_pares_rf()
plot_polar_360_sigma_diff_rf()
plot_polar_360_sigma_ratio_rf()

```

=====

Parte E – MODELO 2: GPR (Gaussian Process)

=====

```

kernel = (
    C(1.0, (1e-3, 1e3))
    * RBF(length_scale=np.ones(X_train_scaled.shape[1]),
          length_scale_bounds=(1e-2, 1e2))

```

```

        + WhiteKernel(noise_level=1.0, noise_level_bounds=(1e-5, 1e1))
    )

gpr = GaussianProcessRegressor(
    kernel=kernel,
    normalize_y=True,
    random_state=42,
    n_restarts_optimizer=10,
)

gpr.fit(X_train_scaled, y_train)

print("\nKernel GPR otimizado:")
print(gpr.kernel_)

metrics_gpr = avaliar_modelo(
    nome="GPR",
    modelo=gpr,
    X_train_s=X_train_scaled,
    X_test_s=X_test_scaled,
    y_train=y_train,
    y_test=y_test,
    nome_fig_base="E_GPR",
)

# Validação cruzada 5-fold - GPR
gpr_cv_model = Pipeline([
    ("scaler", StandardScaler()),
    ("gpr", GaussianProcessRegressor(
        kernel=kernel,
        normalize_y=True,
        random_state=42,
        n_restarts_optimizer=5,
    ))
])

gpr_cv_scores = cross_val_score(
    gpr_cv_model,
    X_cv, y_cv,
    cv=5,
    scoring="r2",
)

print(f"\n[GPR] CV 5-fold -> R² médio = {gpr_cv_scores.mean():.3f} ± {gpr_cv_scores.std():.3f}")

# Teste rápido de diferença treino-teste
diff_r2_gpr = metrics_gpr["r2_train"] - metrics_gpr["r2_test"]

```

```

print(f"\n $\Delta R^2$  (treino - teste) GPR = {diff_r2_gpr:.3f}")

# 4) CURVAS  $\sigma_1 - \sigma_3 \times \beta$  ( $-180^\circ$  a  $+180^\circ$ ) - GPR

def gerar_curvas_e_equacoes_pares_gpr():
    """
    Curvas  $(\sigma_1 - \sigma_3) \times \beta$  em  $\beta \in [-180, 180]$  com período  $180^\circ$ 
    e tabela com equações pares completas (GPR).
    """
    fig, ax = plt.subplots(figsize=(10, 6))
    equations = []

    for i, s3 in enumerate(sigmas3_vals):
        sub = df[df["sigma3"] == s3].copy()
        if sub.empty:
            continue

        cor = COLOR_SIGMA3[i % len(COLOR_SIGMA3)]

        y_real = sub["sigma1"] - s3
        ax.scatter(sub["beta"], y_real, alpha=0.4, s=30, color=cor)

        beta_eff_base = np.linspace(0, 90, 181)

        r_ref = sub["r"].mean()
        sigmac_ref = sub["sigmac"].mean()

        grid_base = pd.DataFrame({
            "beta_eff": beta_eff_base,
            "r": r_ref,
            "sigmac": sigmac_ref,
            "sigma3": s3,
        })

        grid_base_scaled =
scaler.transform(grid_base[feature_cols].values)
        signal_base = gpr.predict(grid_base_scaled)
        y_base = signal_base - s3

        # Ajuste PAR COMPLETO
        z = beta_eff_base**2
        A, B, Cc = np.polyfit(z, y_base, deg=2)
        a4 = A
        a2 = B
        a0 = Cc

        eq_str = (
            f"{a4:.4e} ·  $\beta_{eff}^4$  "

```

```

        f"+ {a2:.4e}·β_eff^2 "
        f"+ {a0:.4e}"
    ).replace("+ -", "- ")
    equations.append((s3, a4, a2, a0, eq_str))

    beta_full = np.linspace(-180, 180, 721)
    beta_eff_full = fold_180_deg(beta_full)

    grid_full = pd.DataFrame({
        "beta_eff": beta_eff_full,
        "r": r_ref,
        "sigmac": sigmac_ref,
        "sigma3": s3,
    })
    grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
    sigma1_full = gpr.predict(grid_full_scaled)
    y_full = sigma1_full - s3

    ax.plot(
        beta_full,
        y_full,
        linewidth=2,
        color=cor,
        label=f"σ3 = {s3} MPa"
    )

    ax.set_xlabel("β (graus)")
    ax.set_ylabel("σ1 - σ3 (MPa)")
    ax.set_title(
        "GPR - Curvas σ1 - σ3 x β (-180° a +180°)\n"
        "Período 180° em β_eff"
    )
    ax.grid(True)
    ax.set_xlim(-180, 180)

    ax.legend(
        title="Nível de σ3",
        loc="upper left",
        bbox_to_anchor=(1.02, 1.0),
        borderaxespad=0.0,
        fontsize=9,
        frameon=True
    )

    fig.tight_layout(rect=[0, 0, 0.80, 1])
    salvar_fig_atual("E_GPR_curvas_sigma_diff_periodo_180")

```

```

eq_df = pd.DataFrame(
    [(s3, a4, a2, a0, eq_str) for (s3, a4, a2, a0, eq_str) in
equations],
    columns=[" $\sigma_3$  (MPa)", "a4", "a2", "a0", "Equação ( $\sigma_1 -$   

 $\sigma_3$ ) ( $\beta_{eff}$ )"]
)
display(style_tabela_cinza(eq_df, "Equações pares - GPR"))

def plot_polar_360_sigma_diff_gpr():
    fig = plt.figure(figsize=(7, 7))
    ax = fig.add_subplot(111, projection="polar")

    fig.subplots_adjust(right=0.80)

    ax.set_theta_zero_location("N")
    ax.set_theta_direction(-1)
    ax.set_thetamin(0)
    ax.set_thetamax(360)
    ax.set_thetagrids(range(0, 361, 30))

    colors = plt.cm.viridis(np.linspace(0, 1, len(sigmas3_vals)))
    handles = []

    for i, s3 in enumerate(sigmas3_vals):
        sub = df[df["sigma3"] == s3].copy()
        if sub.empty:
            continue

        beta_real = sub["beta"].values
        theta_real = np.deg2rad((beta_real + 360) % 360)
        r_real = (sub["sigma1"] - s3).values

        h1 = ax.scatter(theta_real, r_real, s=40,
                        color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
alpha=0.85)
        handles.append((h1, f"Real  $\sigma_3$  = {s3} MPa"))

        beta_full = np.linspace(0, 360, 721)
        beta_eff_full = fold_180_deg(beta_full)

        r_ref = sub["r"].mean()
        sigmac_ref = sub["sigmac"].mean()

        grid_full = pd.DataFrame({
            "beta_eff": beta_eff_full,
            "r": r_ref,
            "sigmac": sigmac_ref,
            "sigma3": s3,

```

```

    })

    grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
    sigma1_pred = gpr.predict(grid_full_scaled)
    r_pred      = sigma1_pred - s3

    theta_full = np.deg2rad(beta_full)
    h2, = ax.plot(theta_full, r_pred,
                    color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
linestyle="--", linewidth=2)
    handles.append((h2, f"Pred  $\sigma_3$  = {s3} MPa (GPR)"))

    ax.set_title(r"( $\sigma_1 - \sigma_3$ ) vs  $\beta$  - GPR (0°-360°, período 180°)",
                  fontsize=13, pad=30)

    ax.grid(alpha=0.3)

    h_objs, labels = zip(*handles)
    ax.legend(h_objs, labels,
              loc="upper left", bbox_to_anchor=(1.02, 1.0),
              borderaxespad=0.0, fontsize=9, frameon=True)

    plt.tight_layout()
    salvar_fig_atual("E_GPR_polar_360_sigma_diff")

def plot_polar_360_sigma_ratio_gpr():
    fig = plt.figure(figsize=(7, 7))
    ax = fig.add_subplot(111, projection="polar")

    fig.subplots_adjust(right=0.80)

    ax.set_theta_zero_location("N")
    ax.set_theta_direction(-1)
    ax.set_thetamin(0)
    ax.set_thetamax(360)
    ax.set_thetagrids(range(0, 361, 30))

    sigmas3_vals_ratio = [s for s in sigmas3_vals if s != 0]

    colors = plt.cm.plasma(np.linspace(0, 1, len(sigmas3_vals_ratio)))
    handles = []

    for i, s3 in enumerate(sigmas3_vals_ratio):
        sub = df[df["sigma3"] == s3].copy()
        if sub.empty:
            continue

```

```

beta_real = sub["beta"].values
theta_real = np.deg2rad((beta_real + 360) % 360)
r_real = (sub["sigma1"] / s3).values

h1 = ax.scatter(theta_real, r_real, s=40,
                color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
alpha=0.85)
handles.append((h1, f"Real  $\sigma_3$  = {s3} MPa"))

beta_full = np.linspace(0, 360, 721)
beta_eff_full = fold_180_deg(beta_full)

r_ref = sub["r"].mean()
sigmac_ref = sub["sigmac"].mean()

grid_full = pd.DataFrame({
    "beta_eff": beta_eff_full,
    "r": r_ref,
    "sigmac": sigmac_ref,
    "sigma3": s3,
})

grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
sigma1_pred = gpr.predict(grid_full_scaled)
r_pred = sigma1_pred / s3

theta_full = np.deg2rad(beta_full)
h2, = ax.plot(theta_full, r_pred,
               color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
linestyle="--", linewidth=2)
handles.append((h2, f"Pred  $\sigma_3$  = {s3} MPa (GPR)"))

ax.set_title(r" $\sigma_1/\sigma_3$  vs  $\beta$  - GPR (0°-360°, período 180°)",
            fontsize=13, pad=30)

ax.grid(alpha=0.3)

h_objs, labels = zip(*handles)
ax.legend(h_objs, labels,
         loc="upper left", bbox_to_anchor=(1.02, 1.0),
         borderaxespad=0.0, fontsize=9, frameon=True)

plt.tight_layout()
salvar_fig_atual("E_GPR_polar_360_sigma_ratio")

# Chamada das funções GPR
gerar_curvas_e_equacoes_pares_gpr()

```

```

plot_polar_360_sigma_diff_gpr()
plot_polar_360_sigma_ratio_gpr()

# =====
# Parte E2 - TESTE DE OVERFITTING NO MODELO GPR (Learning Curve)
# =====

print("\n=====")
print("      TESTE DE OVERFITTING - MODELO GPR (Learning Curve)")
print("=====\\n")

# Usa o MESMO X, y e kernel já definidos ( $\beta_{\text{eff}}$ , r,  $\sigma_{\text{mac}}$ ,  $\sigma_3$ )
X_over = X_cv
y_over = y_cv

gpr_over = GaussianProcessRegressor(
    kernel=kernel,
    normalize_y=True,
    random_state=42,
    n_restarts_optimizer=5
)

gpr_pipeline_over = Pipeline([
    ('scaler', StandardScaler()),
    ('gpr', gpr_over)
])

# Split simples treino/teste - comparação rápida
X_train_over, X_test_over, y_train_over, y_test_over =
train_test_split(
    X_over, y_over,
    test_size=0.20,
    random_state=42,
    stratify=df['sigma3']
)

gpr_pipeline_over.fit(X_train_over, y_train_over)

y_pred_train_over = gpr_pipeline_over.predict(X_train_over)
y_pred_test_over  = gpr_pipeline_over.predict(X_test_over)

r2_train_over    = r2_score(y_train_over, y_pred_train_over)
rmse_train_over  = np.sqrt(mean_squared_error(y_train_over,
y_pred_train_over))

r2_test_over     = r2_score(y_test_over, y_pred_test_over)
rmse_test_over   = np.sqrt(mean_squared_error(y_test_over,
y_pred_test_over))

```

```

print("===== SPLIT ÚNICO (20% TESTE) =====")
print(f"[GPR] Treino -> R² = {r2_train_over:.3f} | RMSE = {rmse_train_over:.3f}")
print(f"[GPR] Teste -> R² = {r2_test_over:.3f} | RMSE = {rmse_test_over:.3f}")

# Curva de aprendizagem (Learning Curve)
cv_over = KFold(n_splits=5, shuffle=True, random_state=42)
train_sizes = np.linspace(0.2, 1.0, 6)

train_sizes_abs, train_scores, val_scores = learning_curve(
    estimator=gpr_pipeline_over,
    X=X_over,
    y=y_over,
    cv=cv_over,
    train_sizes=train_sizes,
    scoring='r2',
    n_jobs=-1,
    shuffle=True
)

train_mean = train_scores.mean(axis=1)
train_std = train_scores.std(axis=1)
val_mean = val_scores.mean(axis=1)
val_std = val_scores.std(axis=1)

plt.figure(figsize=(8, 6))
plt.title("Learning Curve - GPR (R² x tamanho do treino)")
plt.grid(alpha=0.3)

plt.fill_between(train_sizes_abs, train_mean - train_std,
                 train_mean + train_std, alpha=0.15, label='Treino ±1σ')
plt.fill_between(train_sizes_abs, val_mean - val_std,
                 val_mean + val_std, alpha=0.15, label='Validação ±1σ')

plt.plot(train_sizes_abs, train_mean, 'o-', linewidth=2, label='R² Treino')
plt.plot(train_sizes_abs, val_mean, 's-', linewidth=2, label='R² Validação')

plt.xlabel("Número de amostras de treino")
plt.ylabel("R²")
plt.legend()
plt.tight_layout()
salvar_fig_atual("E2_GPR_learning_curve")

```

```

print("\nINTERPRETAÇÃO AUTOMÁTICA:")
print("- Se  $R^2$  de treino  $\approx 1.0$  e  $R^2$  de validação bem mais baixo -> forte overfitting.")
print("- Se as curvas se aproximam conforme aumenta o treino -> modelo mais estável.")
print("- Se ambos  $R^2$  são baixos -> underfitting.")
print("- Se  $R^2$  de validação é próximo do treino -> bom compromisso viés/variação.\n")

```

Parte F – MODELO 3: SVR (SVM REGRESSOR)

```

from sklearn.svm import SVR
from sklearn.model_selection import GridSearchCV, cross_val_score, learning_curve
from sklearn.pipeline import Pipeline
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import r2_score, mean_squared_error
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# -----
# F0) GridSearch para SVR
# -----

param_grid = {
    "C": [30, 50, 100, 200, 300],
    "epsilon": [0.01, 0.1, 0.5, 1.0],
    "gamma": ["scale", 0.1, 0.05, 0.01],
}

svr_base = SVR(kernel="rbf")

grid_svr = GridSearchCV(
    estimator=svr_base,
    param_grid=param_grid,
    cv=5,
    scoring="r2",
    n_jobs=-1,
)

grid_svr.fit(X_train_scaled, y_train)
print("\nMelhores parâmetros SVR:", grid_svr.best_params_)

```

```

best_svr = grid_svr.best_estimator_

# -----
# F1) Avaliação padrão (sua função)
# -----

metrics_svr = avaliar_modelo(
    nome="SVR",
    modelo=best_svr,
    X_train_s=X_train_scaled,
    X_test_s=X_test_scaled,
    y_train=y_train,
    y_test=y_test,
    nome_fig_base="F_SVR",
)

# =====
# F2) TESTE DE OVERFITTING 1 – GAP TREINO vs TESTE
# =====

y_train_pred = best_svr.predict(X_train_scaled)
y_test_pred = best_svr.predict(X_test_scaled)

r2_train = r2_score(y_train, y_train_pred)
r2_test = r2_score(y_test, y_test_pred)
rmse_train = np.sqrt(mean_squared_error(y_train, y_train_pred))
rmse_test = np.sqrt(mean_squared_error(y_test, y_test_pred))

print("\n[SVR] Diagnóstico de overfitting (gap treino vs teste):")
print(f"R² treino = {r2_train:.3f}")
print(f"R² teste = {r2_test:.3f}")
print(f"RMSE treino = {rmse_train:.3f} MPa")
print(f"RMSE teste = {rmse_test:.3f} MPa")
print(f"ΔR² (treino - teste) = {(r2_train - r2_test):.3f}")
print(f"ΔRMSE (teste - treino) = {(rmse_test - rmse_train):.3f} MPa")

# =====
# F3) TESTE DE OVERFITTING 2 – CURVA DE APRENDIZADO (Learning Curve)
# =====
# Aqui usa um pipeline "limpo" (scaler + SVR) para não vazar
informação.
# Importante: learning_curve espera X "cru" (sem scaler aplicado
previamente).

svr_lc_model = Pipeline([
    ("scaler", StandardScaler()),
    ("svr", SVR(
        kernel="rbf",

```

```

        C=grid_svr.best_params_["C"],
        epsilon=grid_svr.best_params_["epsilon"],
        gamma=grid_svr.best_params_["gamma"],
    ))
])

train_sizes, train_scores, val_scores = learning_curve(
    estimator=svr_lc_model,
    X=X_cv,
    y=y_cv,
    cv=5,
    scoring="r2",
    train_sizes=np.linspace(0.1, 1.0, 6),
    n_jobs=-1,
)

train_mean = train_scores.mean(axis=1)
train_std = train_scores.std(axis=1)
val_mean = val_scores.mean(axis=1)
val_std = val_scores.std(axis=1)

plt.figure(figsize=(7, 4))
plt.plot(train_sizes, train_mean, "o-", label="Treino")
plt.plot(train_sizes, val_mean, "o-", label="Validação (CV)")
plt.fill_between(train_sizes, train_mean - train_std, train_mean +
train_std, alpha=0.15)
plt.fill_between(train_sizes, val_mean - val_std, val_mean +
val_std, alpha=0.15)
plt.xlabel("Tamanho do conjunto de treino")
plt.ylabel("R²")
plt.title("SVR - Curva de aprendizado (detecção de overfitting)")
plt.grid(alpha=0.3)
plt.legend()
plt.tight_layout()
salvar_fig_atual("F_SVR_learning_curve")

# -----
# F4) Validação cruzada 5-fold - SVR
# -----

svr_cv_model = Pipeline([
    ("scaler", StandardScaler()),
    ("svr", SVR(
        kernel="rbf",
        C=grid_svr.best_params_["C"],
        epsilon=grid_svr.best_params_["epsilon"],
        gamma=grid_svr.best_params_["gamma"],
    ))
])

```

```

])

svr_cv_scores = cross_val_score(
    svr_cv_model,
    X_cv, y_cv,
    cv=5,
    scoring="r2",
)

print(f"\n[SVR] CV 5-fold -> R² médio = {svr_cv_scores.mean():.3f} ±
{svr_cv_scores.std():.3f}")

# =====
# F5) CURVAS  $\sigma_1 - \sigma_3 \times \beta$  ( $-180^\circ$  a  $+180^\circ$ ) - SVR
# =====

def gerar_curvas_e_equacoes_pares_svr():
    """
    Curvas  $(\sigma_1 - \sigma_3) \times \beta$  em  $\beta \in [-180, 180]$  com período  $180^\circ$ 
    e tabela com equações pares completas (SVR).
    """
    fig, ax = plt.subplots(figsize=(10, 6))
    equations = []

    for i, s3 in enumerate(sigmas3_vals):
        sub = df[df["sigma3"] == s3].copy()
        if sub.empty:
            continue

        cor = COLOR_SIGMA3[i % len(COLOR_SIGMA3)]

        y_real = sub["sigma1"] - s3
        ax.scatter(sub["beta"], y_real, alpha=0.4, s=30, color=cor)

        beta_eff_base = np.linspace(0, 90, 181)

        r_ref = sub["r"].mean()
        sigmac_ref = sub["sigmac"].mean()

        grid_base = pd.DataFrame({
            "beta_eff": beta_eff_base,
            "r": r_ref,
            "sigmac": sigmac_ref,
            "sigma3": s3,
        })

        grid_base_scaled =
scaler.transform(grid_base[feature_cols].values)

```

```

signal_base = best_svr.predict(grid_base_scaled)
y_base = signal_base - s3

# Ajuste PAR COMPLETO
z = beta_eff_base**2
A, B, Cc = np.polyfit(z, y_base, deg=2)
a4 = A
a2 = B
a0 = Cc

eq_str = (
    f"{a4:.4e} ·  $\beta_{\text{eff}}^4$  "
    f"+ {a2:.4e} ·  $\beta_{\text{eff}}^2$  "
    f"+ {a0:.4e}"
).replace("+ -", "- ")
equations.append((s3, a4, a2, a0, eq_str))

beta_full = np.linspace(-180, 180, 721)
beta_eff_full = fold_180_deg(beta_full)

grid_full = pd.DataFrame({
    "beta_eff": beta_eff_full,
    "r": r_ref,
    "sigmac": sigmac_ref,
    "sigma3": s3,
})
grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
signal_full = best_svr.predict(grid_full_scaled)
y_full = signal_full - s3

ax.plot(
    beta_full,
    y_full,
    linewidth=2,
    color=cor,
    label=f" $\sigma_3 = \{s3\}$  MPa"
)

ax.set_xlabel(" $\beta$  (graus)")
ax.set_ylabel(" $\sigma_1 - \sigma_3$  (MPa)")
ax.set_title(
    "SVR - Curvas  $\sigma_1 - \sigma_3$  x  $\beta$  (-180° a +180°)\n"
    "Período 180° em  $\beta_{\text{eff}}$ "
)
ax.grid(True)
ax.set_xlim(-180, 180)

```

```

ax.legend(
    title="Nível de  $\sigma_3$ ",
    loc="upper left",
    bbox_to_anchor=(1.02, 1.0),
    borderaxespad=0.0,
    fontsize=9,
    frameon=True
)

fig.tight_layout(rect=[0, 0, 0.80, 1])
salvar_fig_atual("F_SVR_curvas_sigma_diff_periodo_180")

eq_df = pd.DataFrame(
    [(s3, a4, a2, a0, eq_str) for (s3, a4, a2, a0, eq_str) in
equations],
    columns=[" $\sigma_3$  (MPa)", "a4", "a2", "a0", "Equação ( $\sigma_1 -$   

 $\sigma_3$ ) ( $\beta_{eff}$ )"]
)
display(style_tabela_cinza(eq_df, "Equações pares - SVR"))

def plot_polar_360_sigma_diff_svr():
    fig = plt.figure(figsize=(7, 7))
    ax = fig.add_subplot(111, projection="polar")

    fig.subplots_adjust(right=0.80)

    ax.set_theta_zero_location("N")
    ax.set_theta_direction(-1)
    ax.set_thetamin(0)
    ax.set_thetamax(360)
    ax.set_theta grids(range(0, 361, 30))

    handles = []

    for i, s3 in enumerate(sigmas3_vals):
        sub = df[df["sigma3"] == s3].copy()
        if sub.empty:
            continue

        beta_real = sub["beta"].values
        theta_real = np.deg2rad((beta_real + 360) % 360)
        r_real = (sub["sigma1"] - s3).values

        h1 = ax.scatter(theta_real, r_real, s=40,
                        color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
alpha=0.85)
        handles.append((h1, f"Real  $\sigma_3$  = {s3} MPa"))

```

```

beta_full = np.linspace(0, 360, 721)
beta_eff_full = fold_180_deg(beta_full)

r_ref      = sub["r"].mean()
sigmac_ref = sub["sigmac"].mean()

grid_full = pd.DataFrame({
    "beta_eff": beta_eff_full,
    "r":        r_ref,
    "sigmac":    sigmac_ref,
    "sigma3":    s3,
})

grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
sigma1_pred = best_svr.predict(grid_full_scaled)
r_pred      = sigma1_pred - s3

theta_full = np.deg2rad(beta_full)
h2, = ax.plot(theta_full, r_pred,
               color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
               linestyle="--", linewidth=2)
handles.append((h2, f"Pred  $\sigma_3$  = {s3} MPa (SVR)"))

ax.set_title(r"( $\sigma_1 - \sigma_3$ ) vs  $\beta$  - SVR (0°-360°, período 180°)",
             fontsize=13, pad=30)

ax.grid(alpha=0.3)

h_objs, labels = zip(*handles)
ax.legend(h_objs, labels,
          loc="upper left", bbox_to_anchor=(1.02, 1.0),
          borderaxespad=0.0, fontsize=9, frameon=True)

plt.tight_layout()
salvar_fig_atual("F_SVR_polar_360_sigma_diff")

def plot_polar_360_sigma_ratio_svr():
    fig = plt.figure(figsize=(7, 7))
    ax = fig.add_subplot(111, projection="polar")

    fig.subplots_adjust(right=0.80)

    ax.set_theta_zero_location("N")
    ax.set_theta_direction(-1)
    ax.set_thetamin(0)
    ax.set_thetamax(360)
    ax.set_theta grids(range(0, 361, 30))

```

```

sigmas3_vals_ratio = [s for s in sigmas3_vals if s != 0]

handles = []

for i, s3 in enumerate(sigmas3_vals_ratio):
    sub = df[df["sigma3"] == s3].copy()
    if sub.empty:
        continue

    beta_real = sub["beta"].values
    theta_real = np.deg2rad((beta_real + 360) % 360)
    r_real = (sub["sigma1"] / s3).values

    h1 = ax.scatter(theta_real, r_real, s=40,
                    color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
alpha=0.85)
    handles.append((h1, f"Real  $\sigma_3$  = {s3} MPa"))

    beta_full = np.linspace(0, 360, 721)
    beta_eff_full = fold_180_deg(beta_full)

    r_ref = sub["r"].mean()
    sigmac_ref = sub["sigmac"].mean()

    grid_full = pd.DataFrame({
        "beta_eff": beta_eff_full,
        "r": r_ref,
        "sigmac": sigmac_ref,
        "sigma3": s3,
    })

    grid_full_scaled =
scaler.transform(grid_full[feature_cols].values)
    sigma1_pred = best_svr.predict(grid_full_scaled)
    r_pred = sigma1_pred / s3

    theta_full = np.deg2rad(beta_full)
    h2, = ax.plot(theta_full, r_pred,
                    color=COLOR_SIGMA3[i % len(COLOR_SIGMA3)],
linestyle="--", linewidth=2)
    handles.append((h2, f"Pred  $\sigma_3$  = {s3} MPa (SVR)"))

    ax.set_title(r" $\sigma_1/\sigma_3$  vs  $\beta$  - SVR ( $0^\circ$ - $360^\circ$ , período  $180^\circ$ )",
                  fontsize=13, pad=30)

    ax.grid(alpha=0.3)

```

```

h_objs, labels = zip(*handles)
ax.legend(h_objs, labels,
          loc="upper left", bbox_to_anchor=(1.02, 1.0),
          borderaxespad=0.0, fontsize=9, frameon=True)

plt.tight_layout()
salvar_fig_atual("F_SVR_polar_360_sigma_ratio")

# Chamada das funções SVR
gerar_curvas_e_equacoes_pares_svr()
plot_polar_360_sigma_diff_svr()
plot_polar_360_sigma_ratio_svr()

```

Parte G – ESTATÍSTICA DESCRITIVA

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import norm
from IPython.display import display

# -----
# FUNÇÃO DE ESTILO - CINZA
# -----
def style_tabela_cinza(df_est, caption, decimals=2):
    df_est = df_est.copy()

    # colunas numéricas
    num_cols = df_est.select_dtypes(include=["number"]).columns
    df_est[num_cols] = df_est[num_cols].round(decimals)

    # formatação explícita
    fmt = {}
    for c in num_cols:
        if pd.api.types.is_integer_dtype(df_est[c]):
            fmt[c] = "{:d}"
        else:
            fmt[c] = "{:.{:s}f}"

    return (
        df_est.style
        .format(fmt)
        .set_caption(caption)
    )

```

```

.set_table_styles([
    {"selector": "caption",
     "props": [("caption-side", "top"),
                ("font-weight", "bold"),
                ("font-size", "14px"),
                ("text-align", "left"),
                ("padding", "6px 0")]},

    {"selector": "th",
     "props": [("background-color", "#E6E6E6"),
                ("font-weight", "bold"),
                ("border", "1px solid #C0C0C0"),
                ("text-align", "center"),
                ("padding", "6px")]},

    {"selector": "td",
     "props": [("border", "1px solid #C0C0C0"),
                ("text-align", "center"),
                ("padding", "6px")]},

    {"selector": "tr:nth-child(even) td",
     "props": [("background-color", "#F7F7F7")]}
])
)

# -----
# CHECAGEM DO DATAFRAME
# -----
if "df" not in globals():
    raise NameError("O dataframe 'df' não está definido. Carregue/crie
o df antes de rodar este bloco.")

# -----
# DICIONÁRIOS (PT-BR)
# -----
col_map_pt = {
    "count": "N",
    "mean": "Média",
    "median": "Mediana",
    "mode": "Moda",
    "std": "Desvio-padrão",
    "var": "Variância",
    "cv_%": "CV (%)",
    "min": "Mínimo",
    "25%": "Q1",
    "75%": "Q3",
    "max": "Máximo",
    "skew": "Assimetria",

```

```

        "kurt": "Curtose",
    }

var_labels = {
    "sigma1": " $\sigma_1$  (MPa)",
    "beta": " $\beta$  (graus)",
    "beta_abs": " $|\beta|$  (graus)",
    "beta_eff": " $\beta_{eff}$  (graus, período 180°)",
    "sigma3": " $\sigma_3$  (MPa)",
    "r": "r (-)",
    "sigmac": " $\sigma_c$  (MPa)",
    "coesao": "Coesão (MPa)",
    "atrito": "Ângulo de atrito (°)",
    "sigma_diff": " $\sigma_1 - \sigma_3$  (MPa)",
    "sigma_ratio": " $\sigma_1/\sigma_3$  (-)",
}

# -----
# G0 - variáveis existentes no df
# -----
vars_base = [v for v in [
    "sigma1", "beta", "beta_abs", "beta_eff",
    "sigma3", "r", "sigmac", "coesao", "atrito"
] if v in df.columns]

vars_comb = []
if "sigma_diff" in df.columns:
    vars_comb.append("sigma_diff")
if "sigma_ratio" in df.columns:
    vars_comb.append("sigma_ratio")

vars_all = vars_base + vars_comb

# =====
# G1 - ESTATÍSTICA DESCRITIVA GLOBAL
# =====
if len(vars_all) > 0:
    df_stats = df[vars_all].describe().T

    df_stats["var"] = df[vars_all].var()
    df_stats["cv_%"] = (df_stats["std"] / df_stats["mean"]) * 100
    df_stats["skew"] = df[vars_all].skew()
    df_stats["kurt"] = df[vars_all].kurt()
    df_stats["median"] = df[vars_all].median()
    df_stats["mode"] = df[vars_all].mode().iloc[0]

    df_stats = df_stats[
        ["count", "mean", "median", "mode",

```

```

        "std", "var", "cv_%",
        "min", "25%", "75%", "max",
        "skew", "kurt"]
    ]

    # rótulos PT no índice
    df_stats.index = df_stats.index.to_series().map(lambda x:
var_labels.get(x, x))
    # rótulos PT nas colunas
    df_stats.rename(columns=col_map_pt, inplace=True)

    # remover linhas  $\beta$  e  $|\beta|$ 
    df_stats = df_stats.drop(index=[" $\beta$  (graus)", " $|\beta|$  (graus)"],
errors="ignore")

    # garantir numérico
    df_stats = df_stats.apply(pd.to_numeric, errors="coerce")

    display(style_tabela_cinza(df_stats, "Estatística descritiva
global", decimals=2))

else:
    print("Nenhuma variável encontrada para estatística descritiva
global.")

# =====
# G2 - ESTATÍSTICAS POR NÍVEL DE  $\sigma_3$ 
# =====
if "sigma3" in df.columns:
    group_vars = [v for v in ["sigma1", "sigma_diff", "sigma_ratio"] if
v in df.columns]

    if len(group_vars) > 0:
        stats_group = (
            df.groupby("sigma3")[group_vars]
            .agg(["count", "mean", "median", "std", "var", "min",
"max"])
        )

        # N geral (usa o count do  $\sigma_1$ )
        N_geral = stats_group[("sigma1", "count")].astype(int)

        # remove os counts internos
        stats_group = stats_group.drop(columns="count", level=1)

    estat_map = {
        "mean": "Média",
        "median": "Mediana",

```

```

        "std":      "Desvio-padrão",
        "var":      "Variância",
        "min":      "Mínimo",
        "max":      "Máximo"
    }

    stats_group.columns = pd.MultiIndex.from_tuples(
        [(var_labels.get(v, v), estat_map.get(s, s)) for v, s in
stats_group.columns],
        names=["Variável", "Estatística"]
    )

    stats_group.index.name = " $\sigma_3$  (MPa)"

    # insere N geral como primeira coluna
    stats_group.insert(0, ("N", ""), N_geral)

    # garantir numérico fora do N
    for c in stats_group.columns:
        if c[0] != "N":
            stats_group[c] = pd.to_numeric(stats_group[c],
errors="coerce")

    display(
        style_tabela_cinza(
            stats_group,
            "Estatísticas por nível de  $\sigma_3$  - Variáveis de
resistência",
            decimals=2
        )
    )
    else:
        print("Nenhuma variável disponível para estatísticas por nível
de  $\sigma_3$ .")
    else:
        print("A coluna 'sigma3' não existe no dataframe.")
    # -----
    # G3 - matriz de correlação
    # -----
    vars_corr = [v for v in ["sigma1", "sigma3", "beta", "r"] if v in
df.columns]

    if len(vars_corr) > 0:

        corr_mat = df[vars_corr].corr()

        plt.figure(figsize=(8, 6))
        ax = sns.heatmap(

```

```

        corr_mat,
        cmap="coolwarm",
        annot=True,
        fmt=".2f",
        linewidths=0.5,
        square=True
    )

    ax.set_xticklabels(
        [var_labels.get(v, v) for v in corr_mat.columns],
        rotation=45, ha="right"
    )
    ax.set_yticklabels(
        [var_labels.get(v, v) for v in corr_mat.index],
        rotation=0
    )

    plt.title("Matriz de correlação - variáveis principais")
    plt.tight_layout()
    salvar_fig_atual("G3_matriz_correlacao")

# -----
# G5 - histogramas + curva Normal
# -----
vars_hist = [v for v in ["sigma1", "beta", "sigma3", "r"] if v in
df.columns]

if len(vars_hist) > 0:

    fig, axes = plt.subplots(2, 3, figsize=(12, 6))
    axes = axes.flatten()

    for i, col in enumerate(vars_hist):
        ax = axes[i]
        data = df[col].dropna()

        mu = data.mean()
        sigma = data.std(ddof=1)

        ax.hist(data, bins=10, density=True, edgecolor="black",
alpha=0.7)

        if sigma > 0:
            x = np.linspace(data.min(), data.max(), 200)
            ax.plot(x, norm.pdf(x, mu, sigma), linewidth=2)

        ax.set_title(var_labels.get(col, col))
        ax.grid(alpha=0.3)

```

```

for j in range(i + 1, len(axes)):
    fig.delaxes(axes[j])

plt.suptitle("Histogramas + curva Normal - variáveis base", y=0.97)
plt.tight_layout(rect=[0, 0.03, 1, 0.95])
salvar_fig_atual("G5_histogramas_base")

```

Parte H – COMPARAÇÃO FINAL DOS MODELOS

```

comparacao = pd.DataFrame([
    {
        "Modelo": "Random Forest",
        "R2_treino": metrics_rf["r2_train"],
        "R2_teste": metrics_rf["r2_test"],
        "RMSE_treino": metrics_rf["rmse_train"],
        "RMSE_teste": metrics_rf["rmse_test"],
        "MAE_treino": metrics_rf["mae_train"],
        "MAE_teste": metrics_rf["mae_test"],
        "R2_CV_medio": rf_cv_scores.mean(),
        "R2_CV_std": rf_cv_scores.std(),
    },
    {
        "Modelo": "GPR",
        "R2_treino": metrics_gpr["r2_train"],
        "R2_teste": metrics_gpr["r2_test"],
        "RMSE_treino": metrics_gpr["rmse_train"],
        "RMSE_teste": metrics_gpr["rmse_test"],
        "MAE_treino": metrics_gpr["mae_train"],
        "MAE_teste": metrics_gpr["mae_test"],
        "R2_CV_medio": gpr_cv_scores.mean(),
        "R2_CV_std": gpr_cv_scores.std(),
    },
    {
        "Modelo": "SVR",
        "R2_treino": metrics_svr["r2_train"],
        "R2_teste": metrics_svr["r2_test"],
        "RMSE_treino": metrics_svr["rmse_train"],
        "RMSE_teste": metrics_svr["rmse_test"],
        "MAE_treino": metrics_svr["mae_train"],
        "MAE_teste": metrics_svr["mae_test"],
        "R2_CV_medio": svr_cv_scores.mean(),
        "R2_CV_std": svr_cv_scores.std(),
    },
],

```

```

))

display(style_tabela_cinza(
    comparacao,
    "Métricas globais dos modelos (incluindo validação cruzada)"
))

# Gráficos de dispersão - comparação visual no teste
y_test_array = np.asarray(y_test)
y_pred_rf = rf.predict(X_test_scaled)
y_pred_gpr = gpr.predict(X_test_scaled)
y_pred_svr = best_svr.predict(X_test_scaled)

fig, axes = plt.subplots(1, 3, figsize=(15, 4), sharex=True,
sharey=True)

model_names = ["Random Forest", "GPR", "SVR"]
pred_list = [y_pred_rf, y_pred_gpr, y_pred_svr]

min_val = min(y_test_array.min(), y_pred_rf.min(), y_pred_gpr.min(),
y_pred_svr.min())
max_val = max(y_test_array.max(), y_pred_rf.max(), y_pred_gpr.max(),
y_pred_svr.max())
pad = 0.05 * (max_val - min_val)
min_lim, max_lim = min_val - pad, max_val + pad

for ax, nome, y_pred in zip(axes, model_names, pred_list):
    ax.scatter(y_test_array, y_pred, alpha=0.7)
    ax.plot([min_lim, max_lim], [min_lim, max_lim], "r--", linewidth=2)
    ax.set_xlim(min_lim, max_lim)
    ax.set_ylim(min_lim, max_lim)
    ax.set_title(nome, fontsize=10)
    ax.set_xlabel(" $\sigma_1$  real (teste)")
    ax.grid(alpha=0.3)

axes[0].set_ylabel(" $\sigma_1$  predito (teste)")
plt.suptitle("Comparação visual -  $\sigma_1$  real vs predito (teste)",
fontsize=13)
plt.tight_layout(rect=[0, 0.03, 1, 0.92])
salvar_fig_atual("H_scatter_RF_GPR_SVR")

# Gráficos de barra - R2 e RMSE
fig, axes = plt.subplots(1, 2, figsize=(12, 4))

sns.barplot(data=comparacao, x="Modelo", y="R2_teste", ax=axes[0])
axes[0].set_title("R2 (teste) entre modelos")
axes[0].set_ylim(0, 1.05)
axes[0].grid(axis="y", alpha=0.3)

```

```

axes[0].set_xlabel("")
axes[0].set_ylabel("R² (teste)")

sns.barplot(data=comparacao, x="Modelo", y="RMSE_teste", ax=axes[1])
axes[1].set_title("RMSE (teste) entre modelos")
axes[1].grid(axis="y", alpha=0.3)
axes[1].set_xlabel("")
axes[1].set_ylabel("RMSE (MPa)")

plt.tight_layout()
salvar_fig_atual("H_barplots_R2_RMSE")

```

--- Parte I – Análises adicionais ---

```

import scipy.stats as st
from matplotlib.backends.backend_pdf import PdfPages

from sklearn.pipeline import Pipeline
from sklearn.model_selection import KFold, cross_val_score
from sklearn.preprocessing import StandardScaler
from sklearn.svm import SVR

blue = "#1f77b4"

# -----
# I0) Preparar dataframe do teste com previsões e resíduos
# -----
# Garante que y_test vire array 1D
y_test_array = np.asarray(y_test).ravel()

# Previsões no conjunto de teste (usando os modelos já treinados)
y_pred_rf = rf.predict(X_test_scaled)
y_pred_gpr = gpr.predict(X_test_scaled)
y_pred_svr = best_svr.predict(X_test_scaled)

# Montar o dataframe do teste
test_df = pd.DataFrame({
    "sigma_real": y_test_array,
    "sigma_pred_rf": y_pred_rf,
    "sigma_pred_gpr": y_pred_gpr,
    "sigma_pred_svr": y_pred_svr,
})

# Resíduos (real - predito)

```

```

test_df["resid_rf"] = test_df["sigmal_real"] -
test_df["sigmal_pred_rf"]
test_df["resid_gpr"] = test_df["sigmal_real"] -
test_df["sigmal_pred_gpr"]
test_df["resid_svr"] = test_df["sigmal_real"] -
test_df["sigmal_pred_svr"]

# -----
# I1) QQ-PLOTS dos resíduos dos três modelos (global)
# -----

fig, axes = plt.subplots(1, 3, figsize=(15, 4))

modelos_res = [
    ("Random Forest", test_df["resid_rf"]),
    ("GPR", test_df["resid_gpr"]),
    ("SVR", test_df["resid_svr"]),
]

for ax, (nome, resid) in zip(axes, modelos_res):
    st.probplot(resid, dist="norm", plot=ax)
    ax.set_title(f"QQ-Plot - {nome}")
    ax.grid(alpha=0.3)

plt.tight_layout()
salvar_fig_atual("K1_QQplots_residuos_global")

# -----
# I2) Histogramas dos resíduos (3 modelos) - global
# -----

plt.figure(figsize=(12, 4))

for i, (nome, resid) in enumerate(modelos_res):
    plt.subplot(1, 3, i + 1)
    plt.hist(
        resid,
        bins=15,
        alpha=0.8,
        edgecolor="black",
        color=blue,
    )
    plt.title(f"Histograma dos resíduos - {nome}")
    plt.xlabel("Resíduo (MPa)")
    plt.ylabel("Frequência")
    plt.grid(alpha=0.3)

plt.tight_layout()

```

```

salvar_fig_atual("K2_hist_residuos_global")

# -----
# I3) Resíduos vs valor predito (heteroscedasticidade)
# -----

fig, axes = plt.subplots(1, 3, figsize=(15, 4))

preds = [
    ("Random Forest", test_df["sigma1_pred_rf"], test_df["resid_rf"]),
    ("GPR", test_df["sigma1_pred_gpr"],
test_df["resid_gpr"]),
    ("SVR", test_df["sigma1_pred_svr"],
test_df["resid_svr"]),
]

for ax, (nome, pred, resid) in zip(axes, preds):
    ax.scatter(pred, resid, alpha=0.6, color=blue)
    ax.axhline(0, color="red", linestyle="--")
    ax.set_title(f"{nome} - Resíduo vs predição")
    ax.set_xlabel("σ1 predito (MPa)")
    ax.set_ylabel("Resíduo (MPa)")
    ax.grid(alpha=0.3)

plt.tight_layout()
salvar_fig_atual("K3_resid_vs_pred")

# -----
# I4) Validação Cruzada (5-fold) para o SVR
# -----

# Usa o banco completo, sem vazios
X_cv_k4 = df[feature_cols].values
y_cv_k4 = df["sigma1"].values

# Extrair parâmetros do best_svr com segurança
svr_params = best_svr.get_params()

svr_cv_model = Pipeline([
    ("scaler", StandardScaler()),
    ("svr", SVR(**svr_params)),
])

cv_k4 = KFold(n_splits=5, shuffle=True, random_state=42)
scores_k4 = cross_val_score(
    svr_cv_model,
    X_cv_k4, y_cv_k4,
    cv=cv_k4,

```

```

        scoring="r2",
    )

print("\n[SVR] Validação Cruzada 5-fold (reavaliação):")
print(f"R² médio = {scores_k4.mean():.3f} ± {scores_k4.std():.3f}")

plt.figure(figsize=(6, 4))
plt.hist(scores_k4, bins=5, edgecolor="black", alpha=0.8, color=blue)
plt.xlabel("R² (fold)")
plt.ylabel("Frequência")
plt.title("Distribuição dos R² - SVR (Validação Cruzada 5-fold)")
plt.grid(alpha=0.3)
plt.tight_layout()
salvar_fig_atual("K4_hist_R2_SVR_CV")

# -----
# I5) Exportar resultados para Excel (com openpyxl) e PDF
# -----

excel_path = BASE_FIG_DIR / "resultados_analise_sigma1.xlsx"

with pd.ExcelWriter(excel_path, engine="openpyxl") as writer:
    if "df_stats" in globals():
        df_stats.to_excel(writer, sheet_name="estat_global")
    if "stats_group" in globals():
        stats_group.to_excel(writer, sheet_name="estat_por_sigma3")
    if "comparacao" in globals():
        comparacao.to_excel(writer, sheet_name="metricas_modelos",
index=False)
    if "metrics_sigma3" in globals():
        metrics_sigma3.to_excel(writer,
sheet_name="metricas_por_sigma3", index=False)

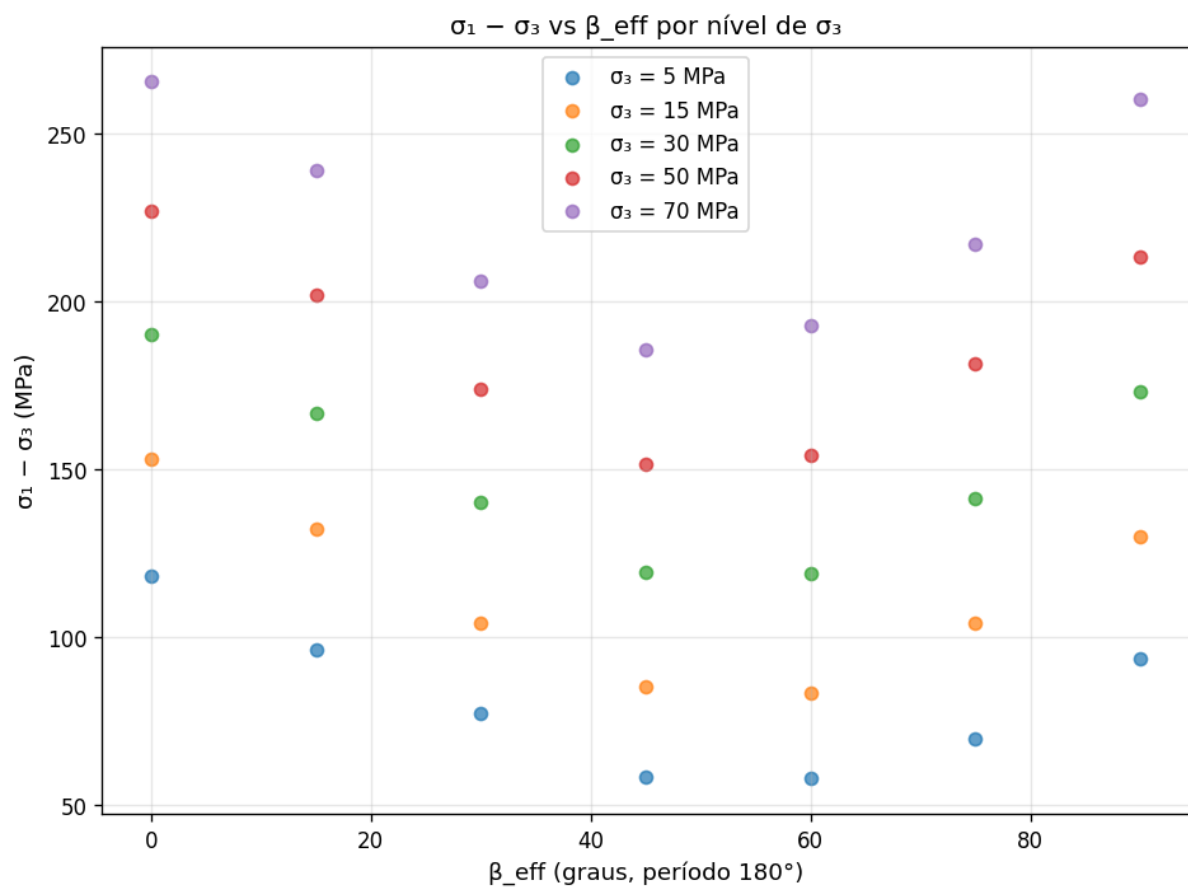
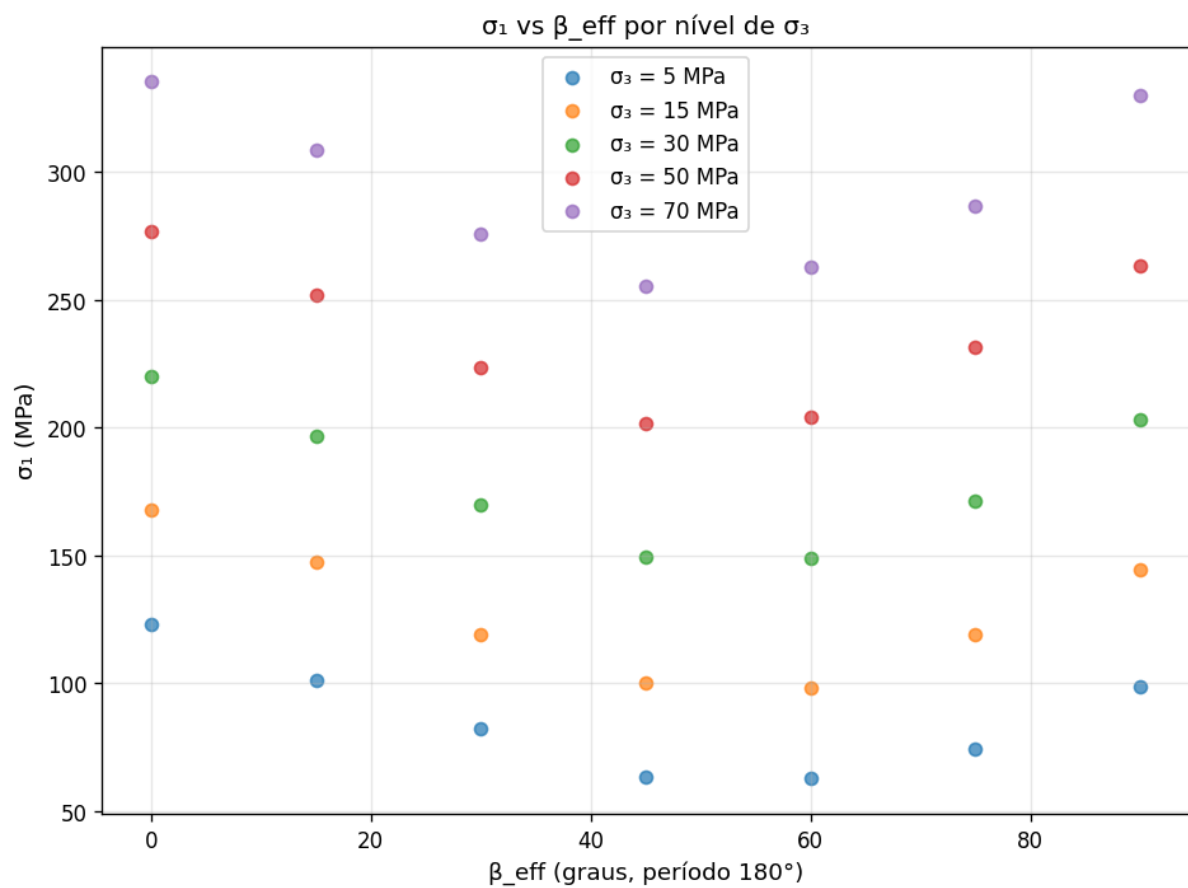
print("\nArquivo Excel gerado com sucesso em:")
print(f" {excel_path}")

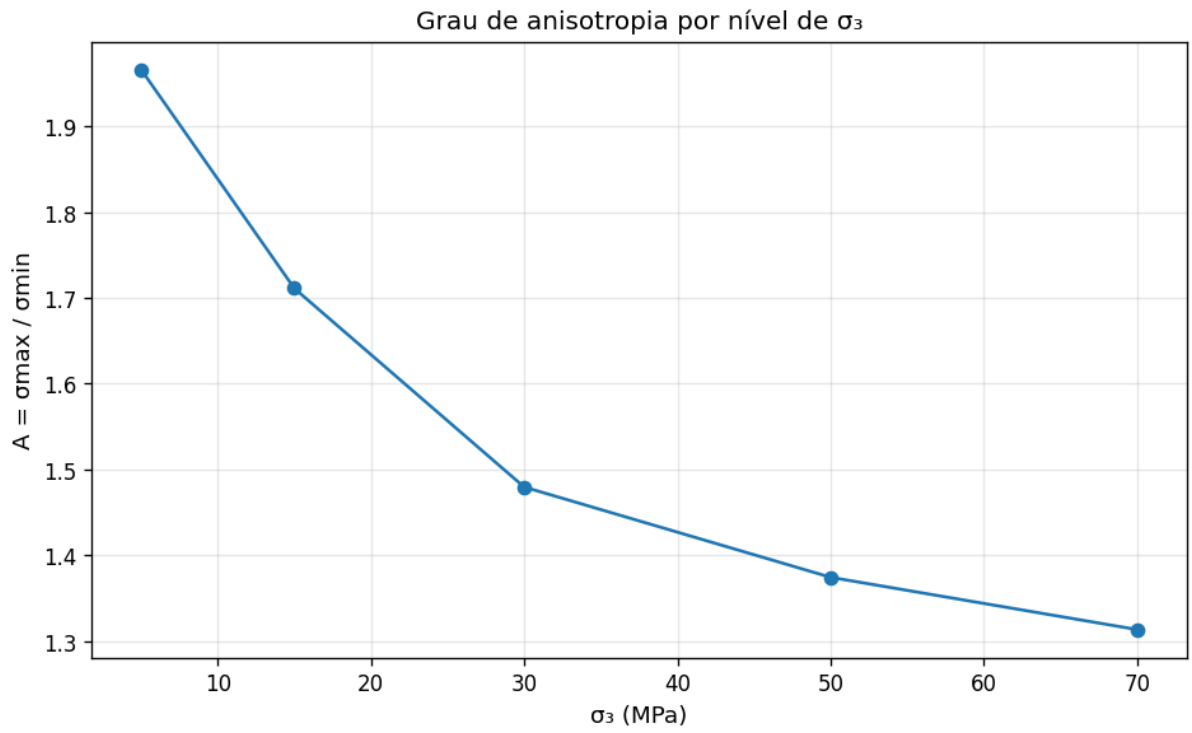
```

2. Resultados do código:

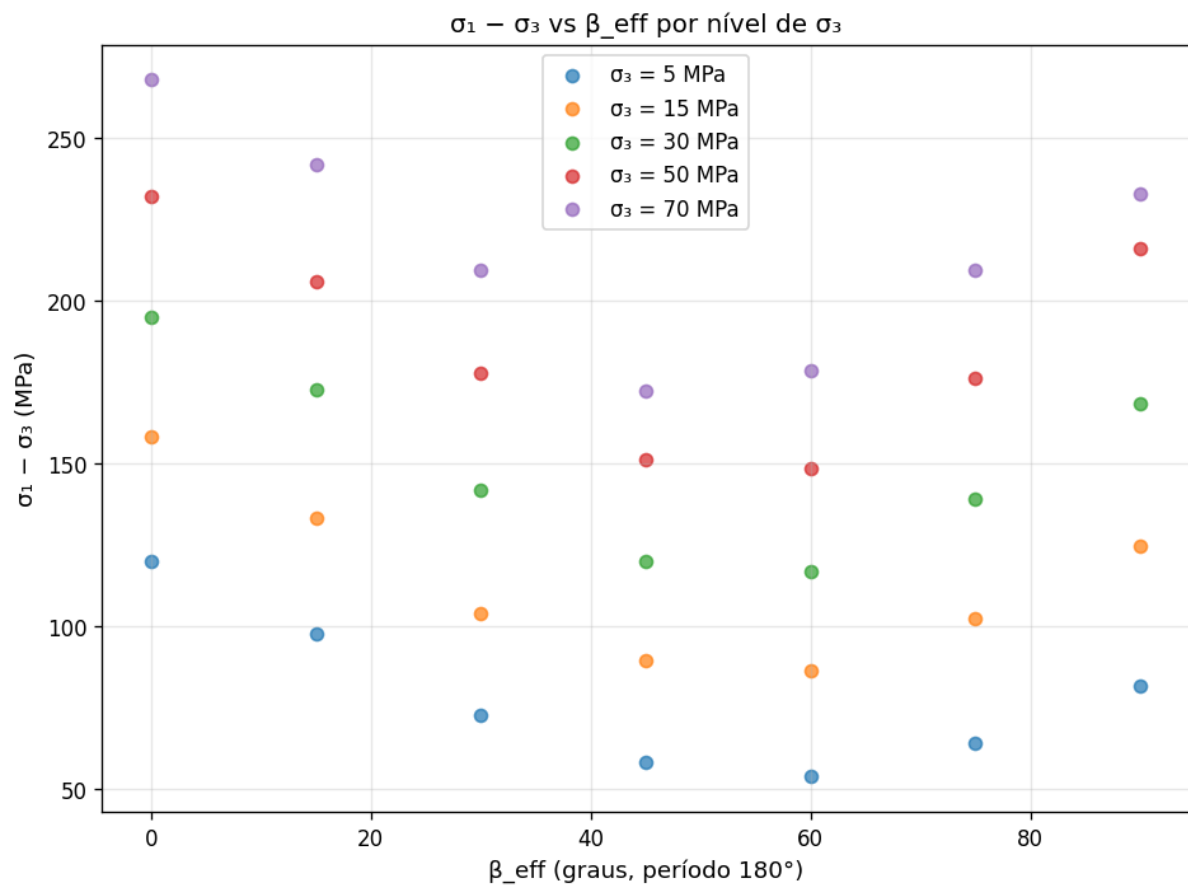
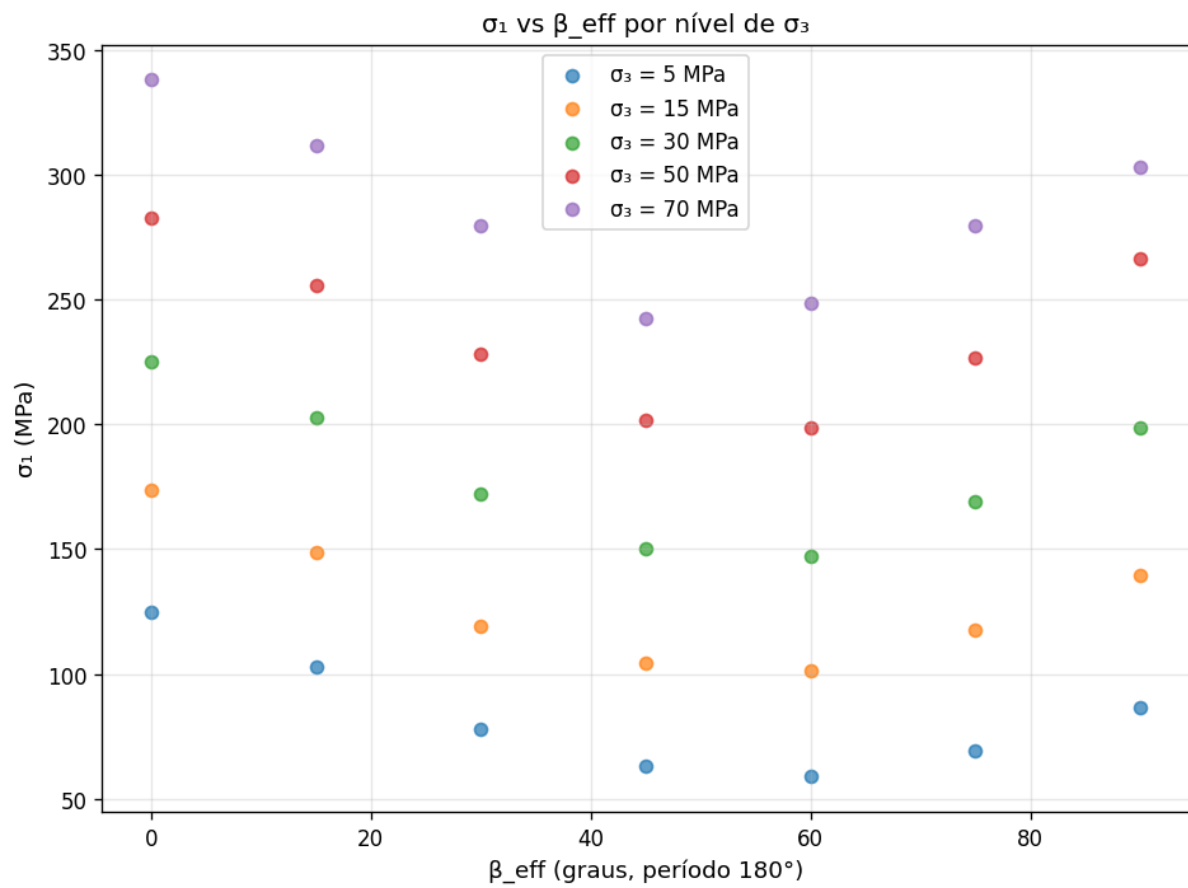
2.1. Análise da anisotropia:

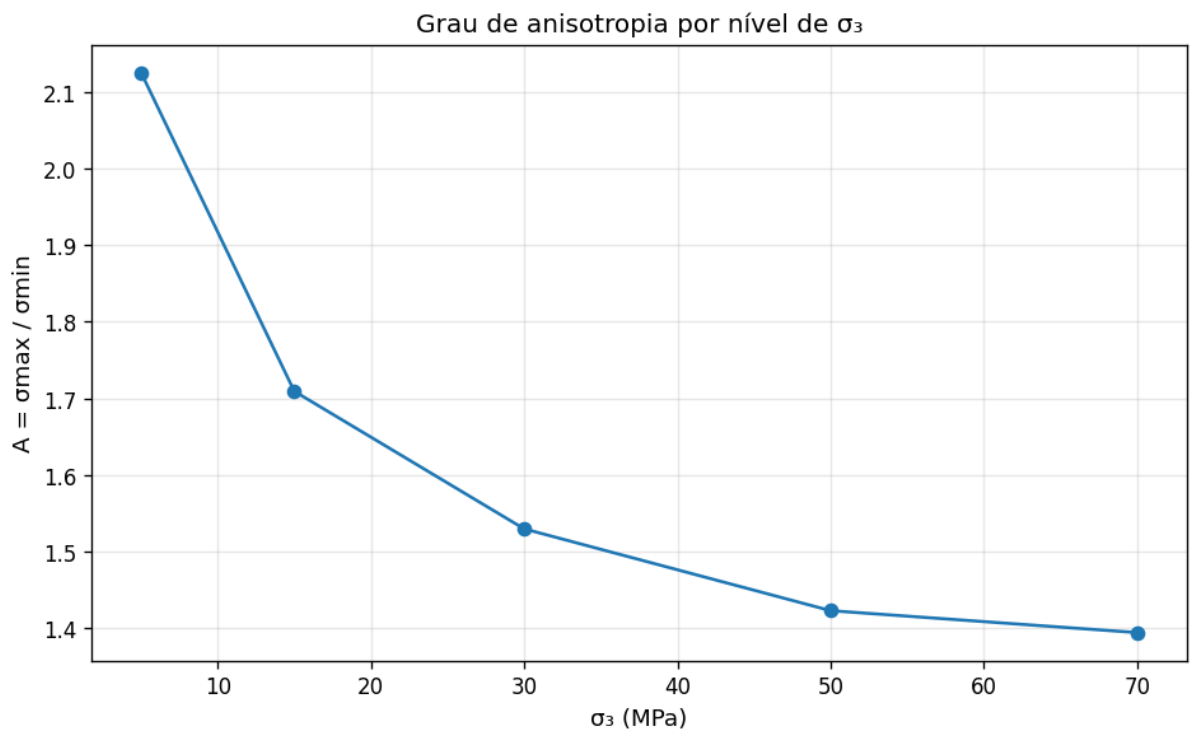
- *Carbonaceous*:



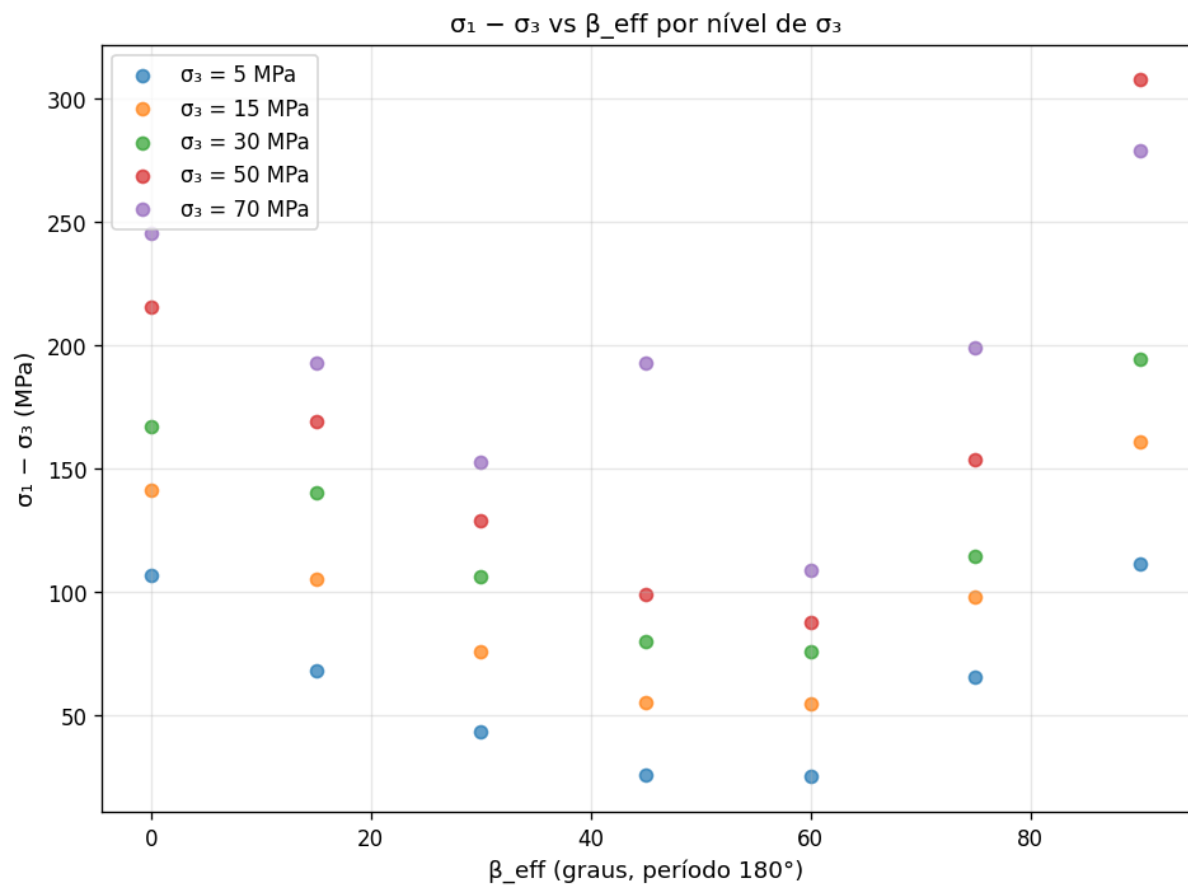
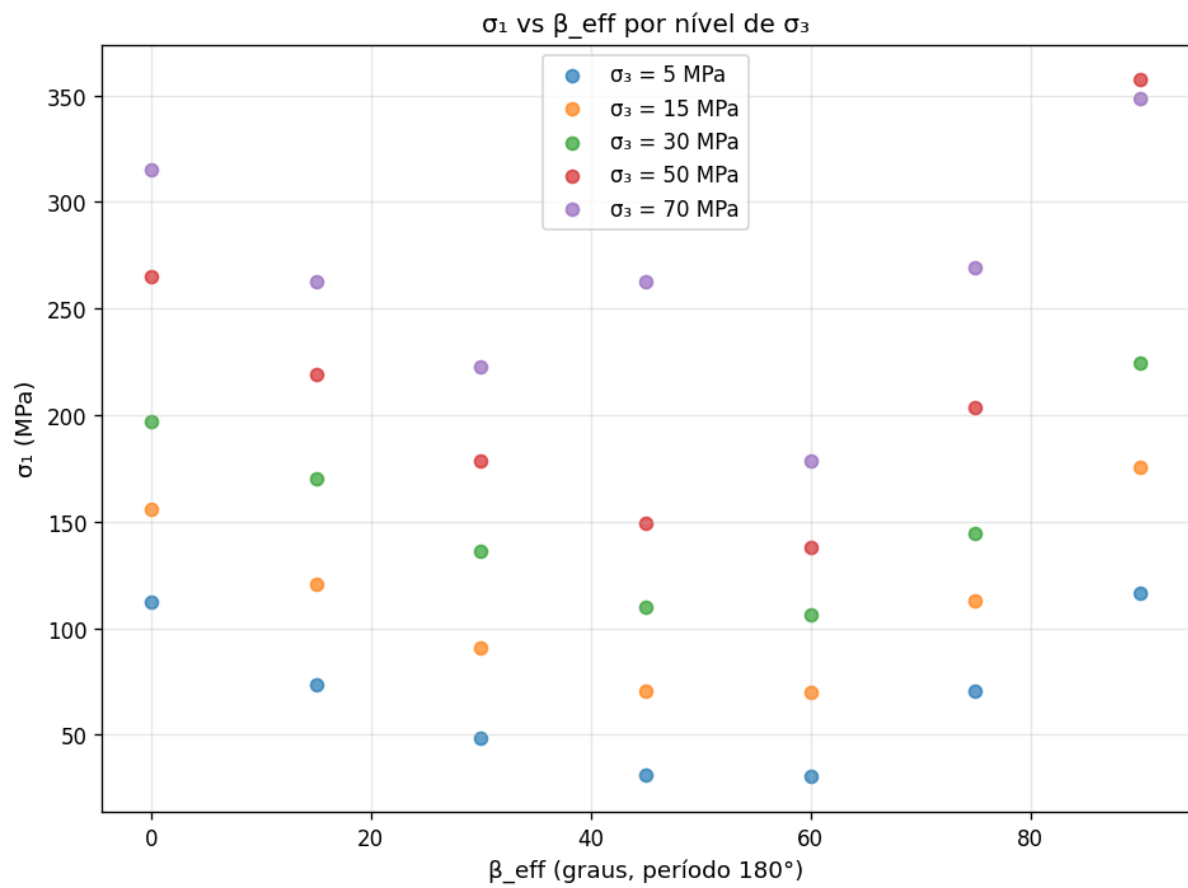


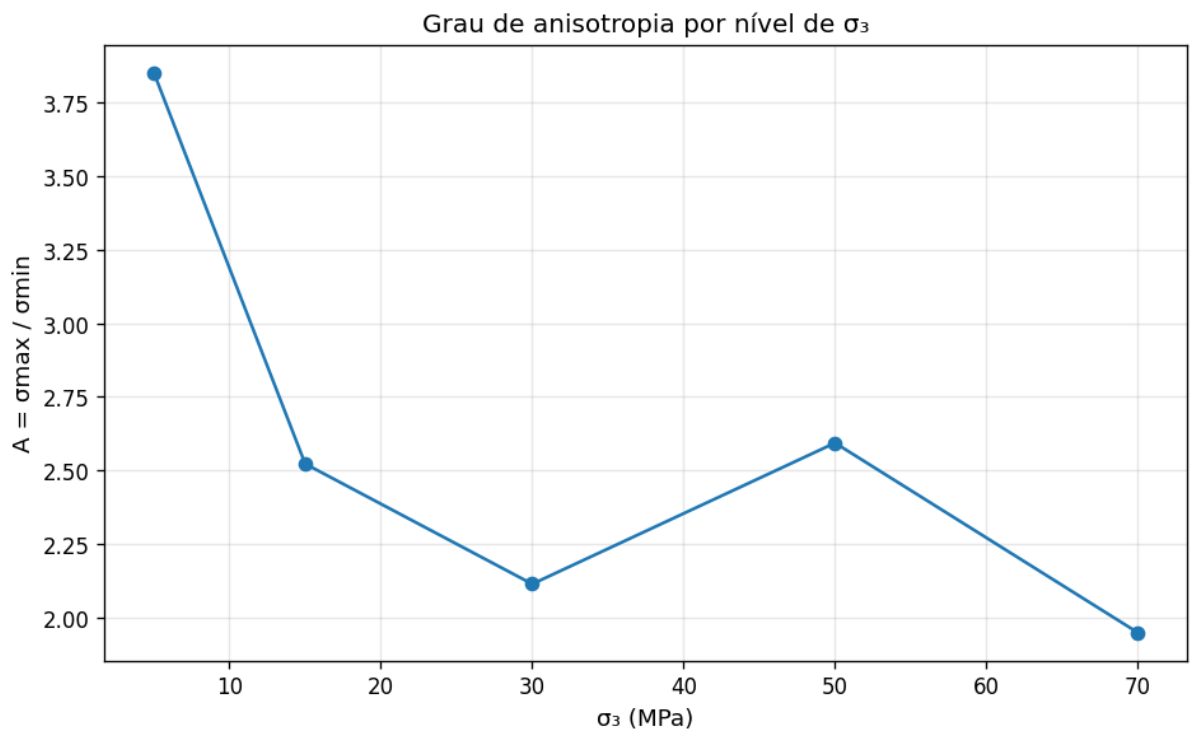
- *Quartzite:*



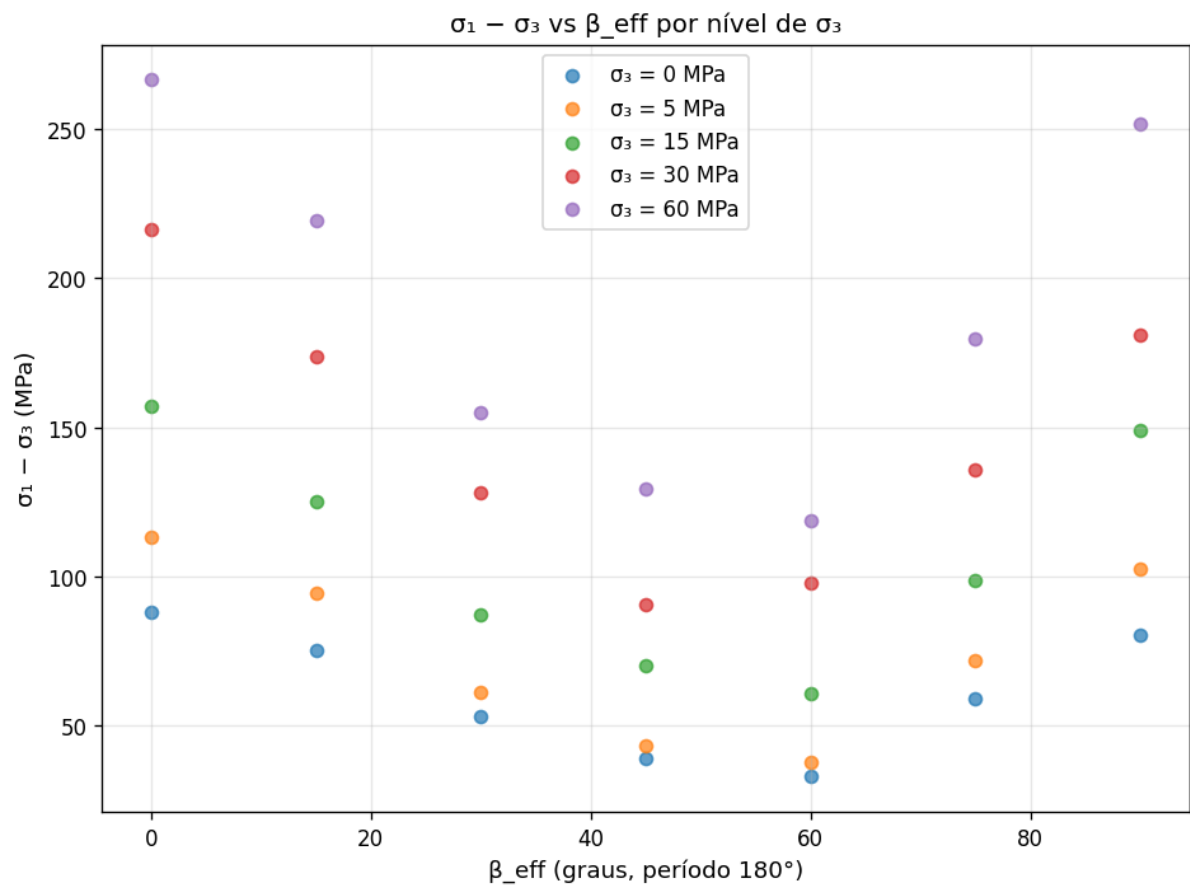
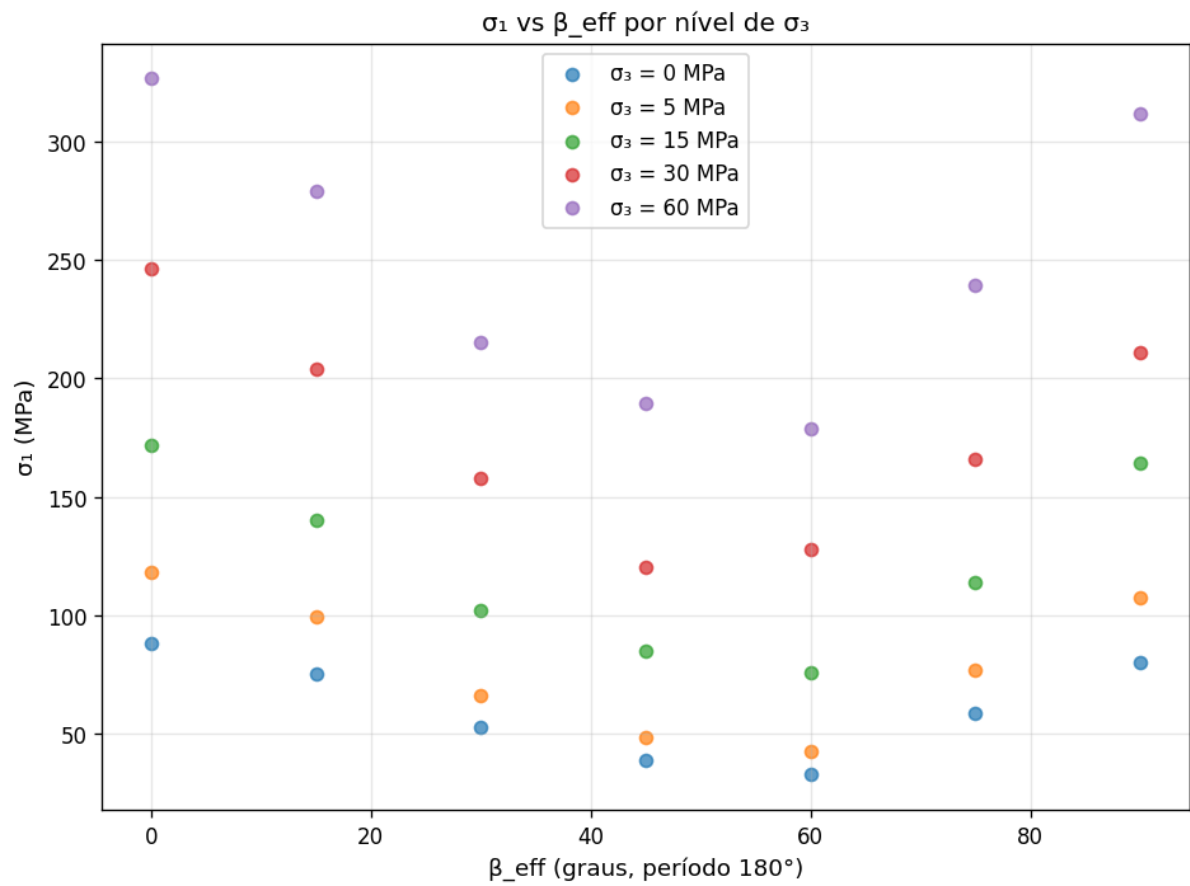


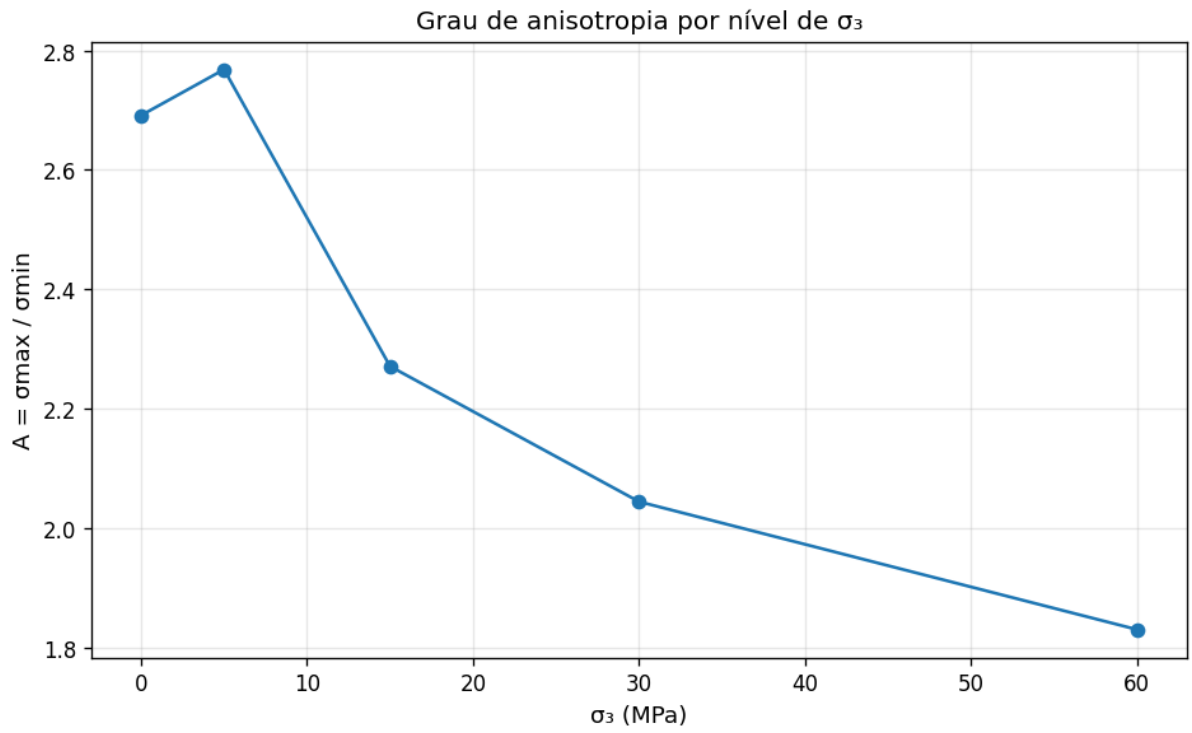
- *Micaceous:*



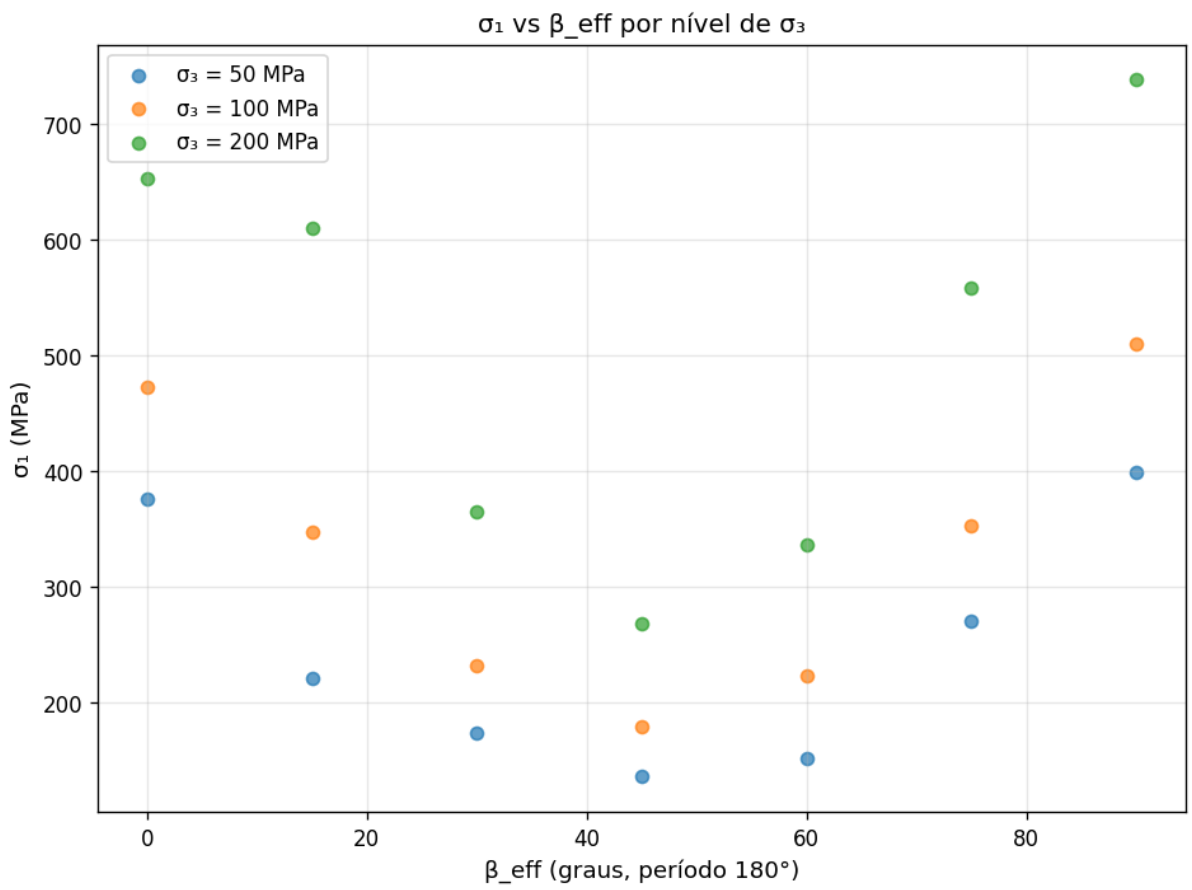


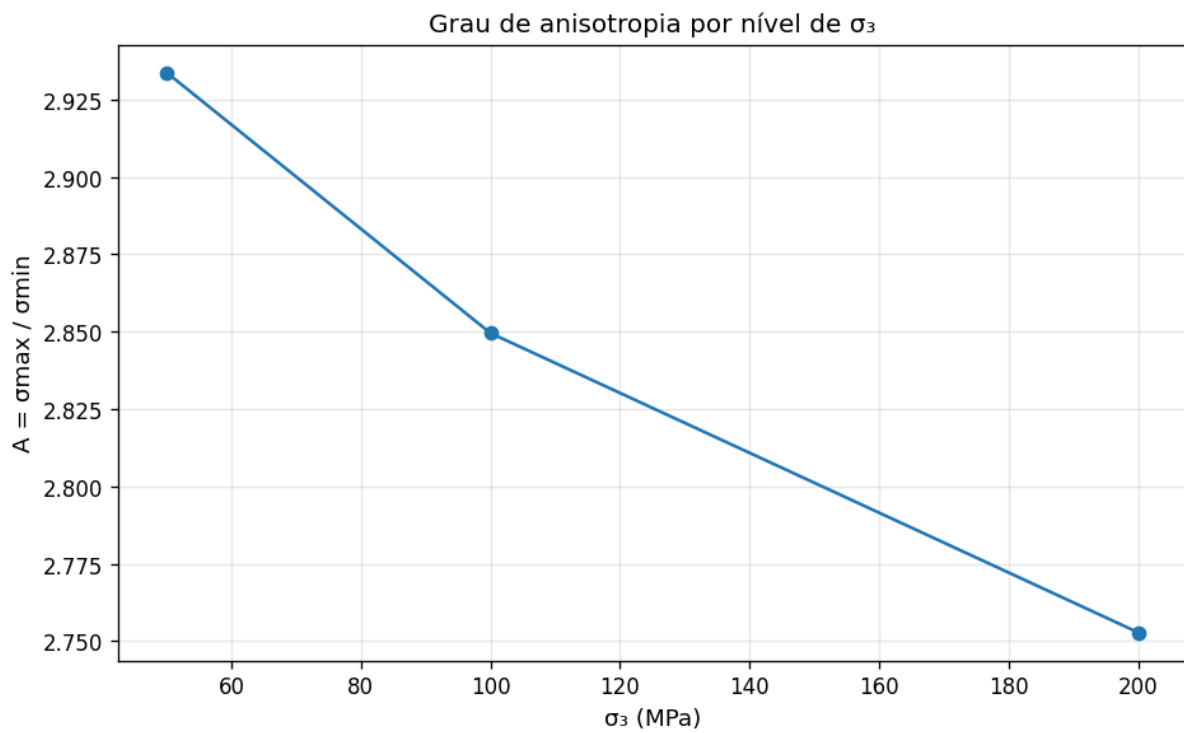
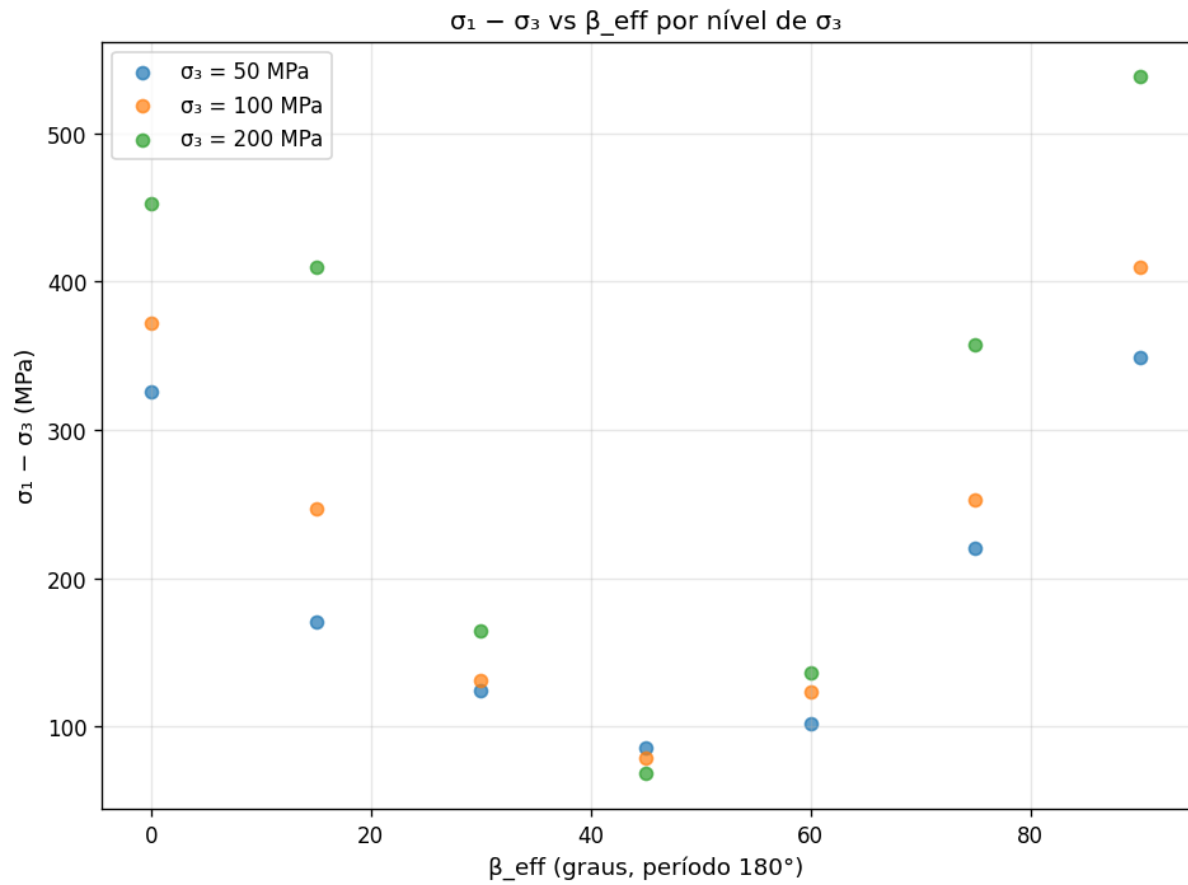
- *Dry:*





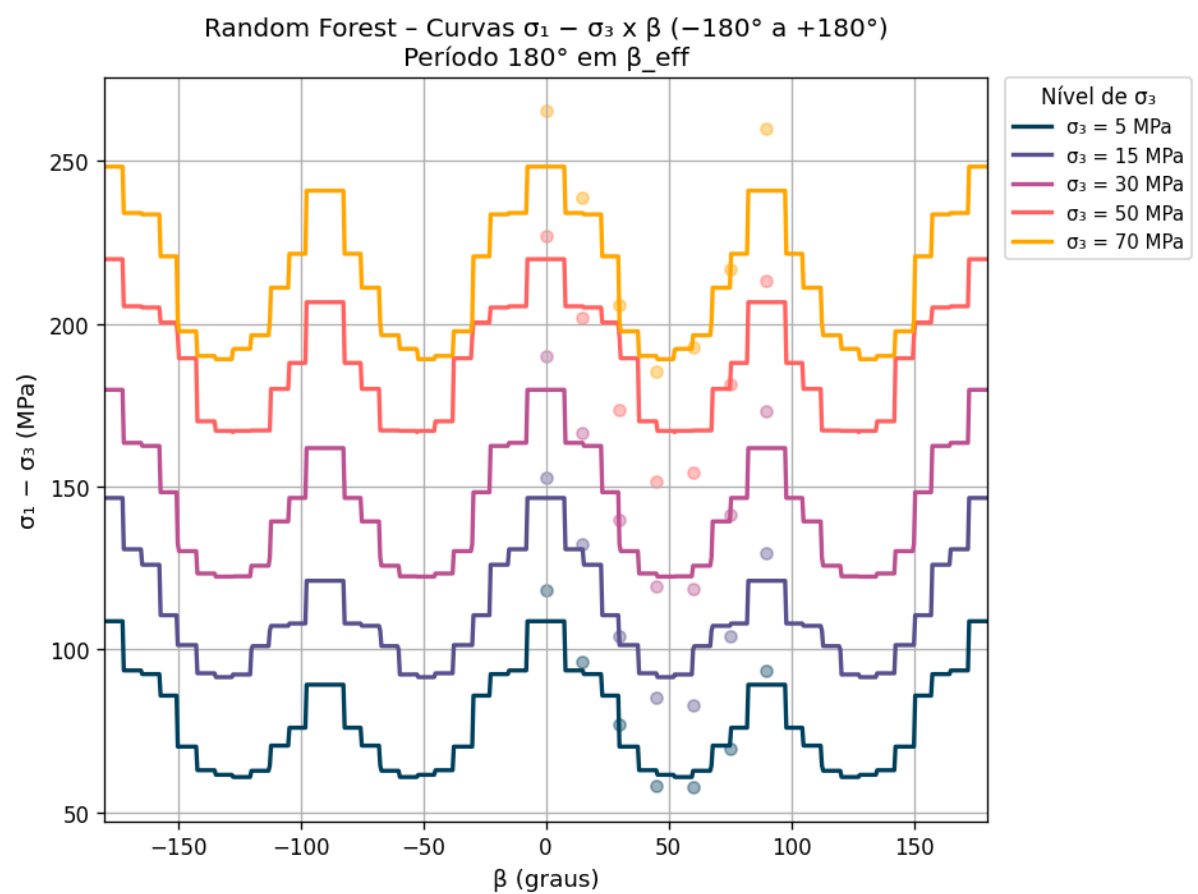
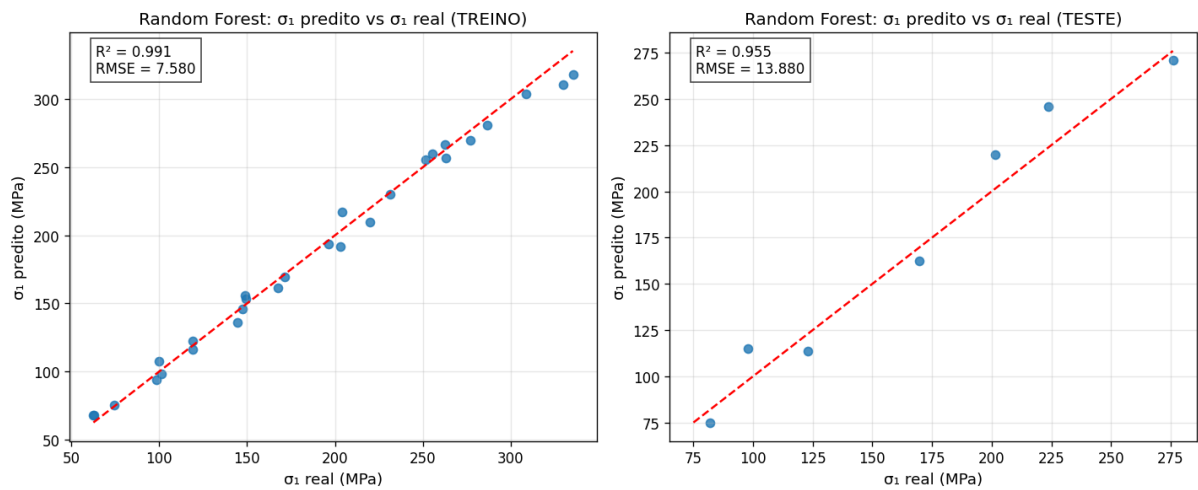
- *Moretown:*



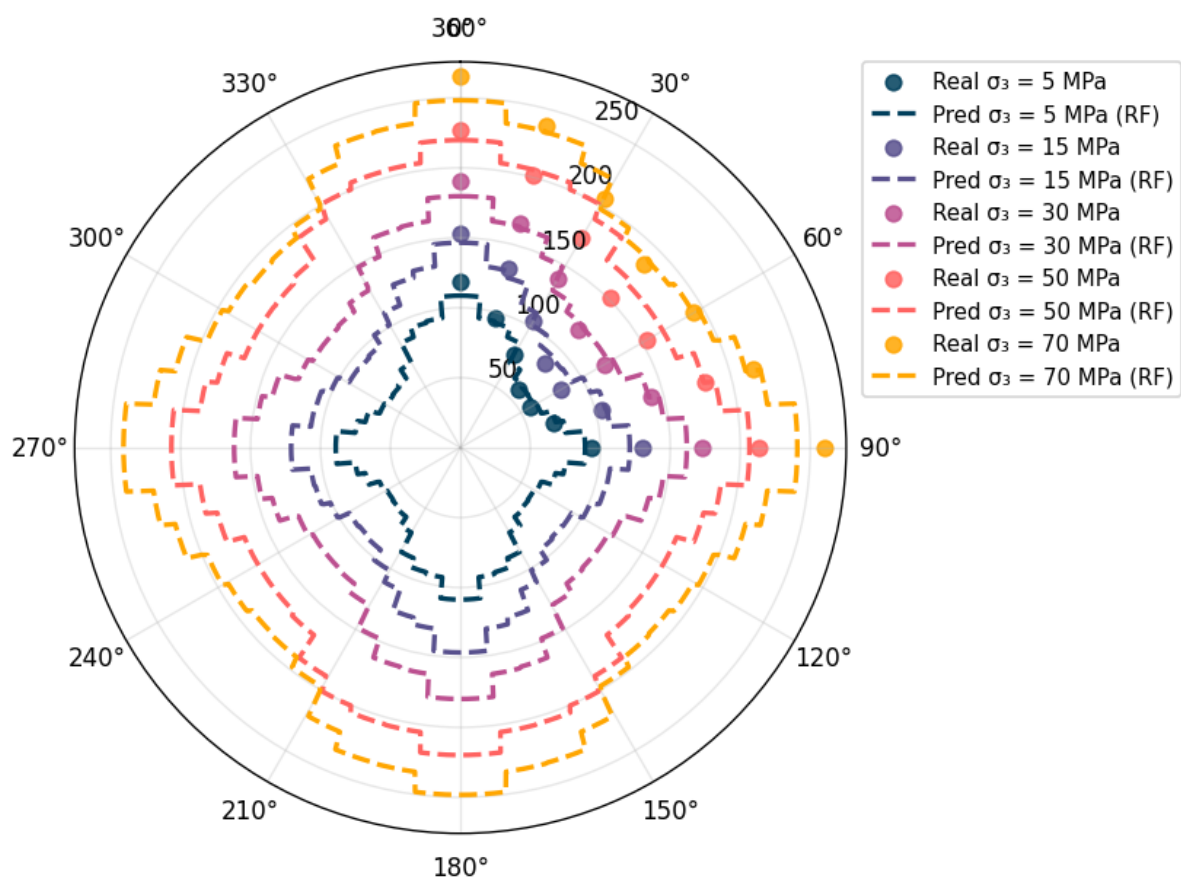


2.2. Resultados do Random Forest:

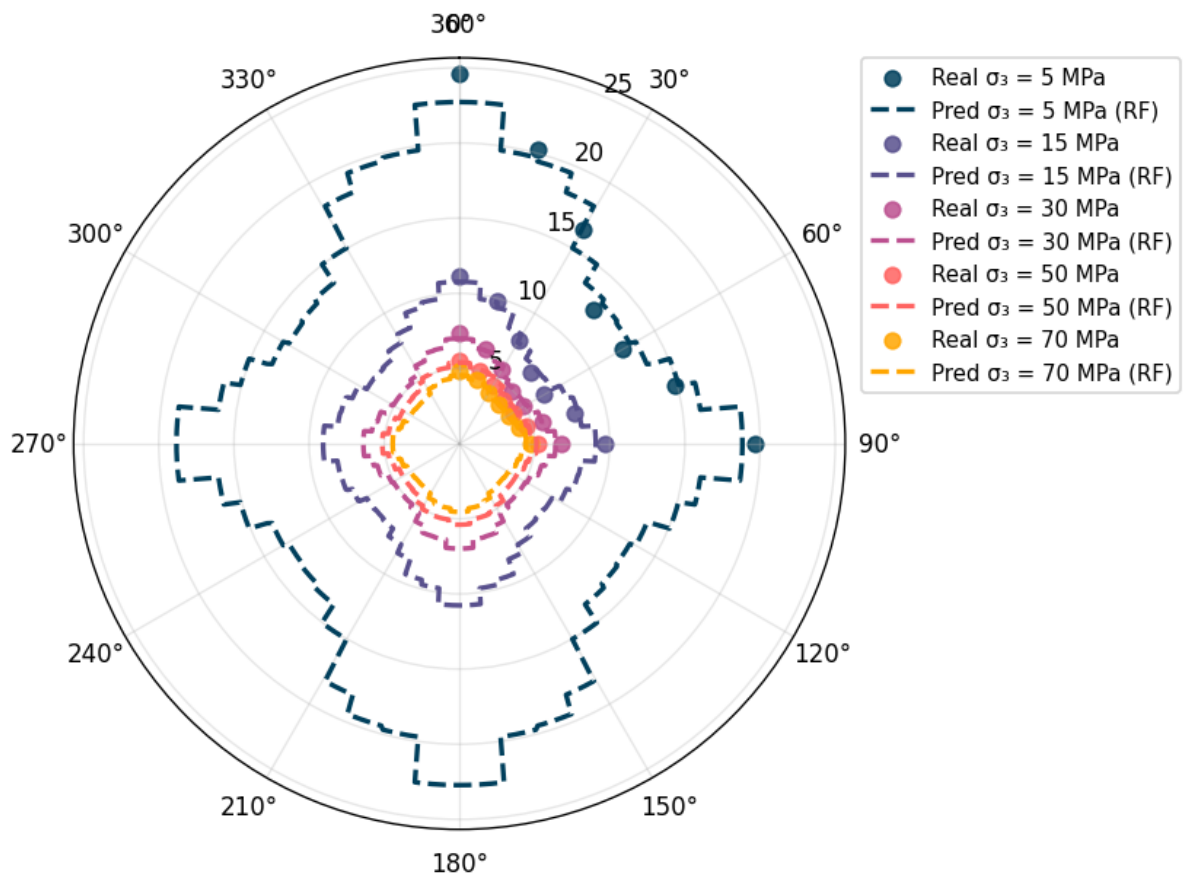
- *Carbonaceous:*



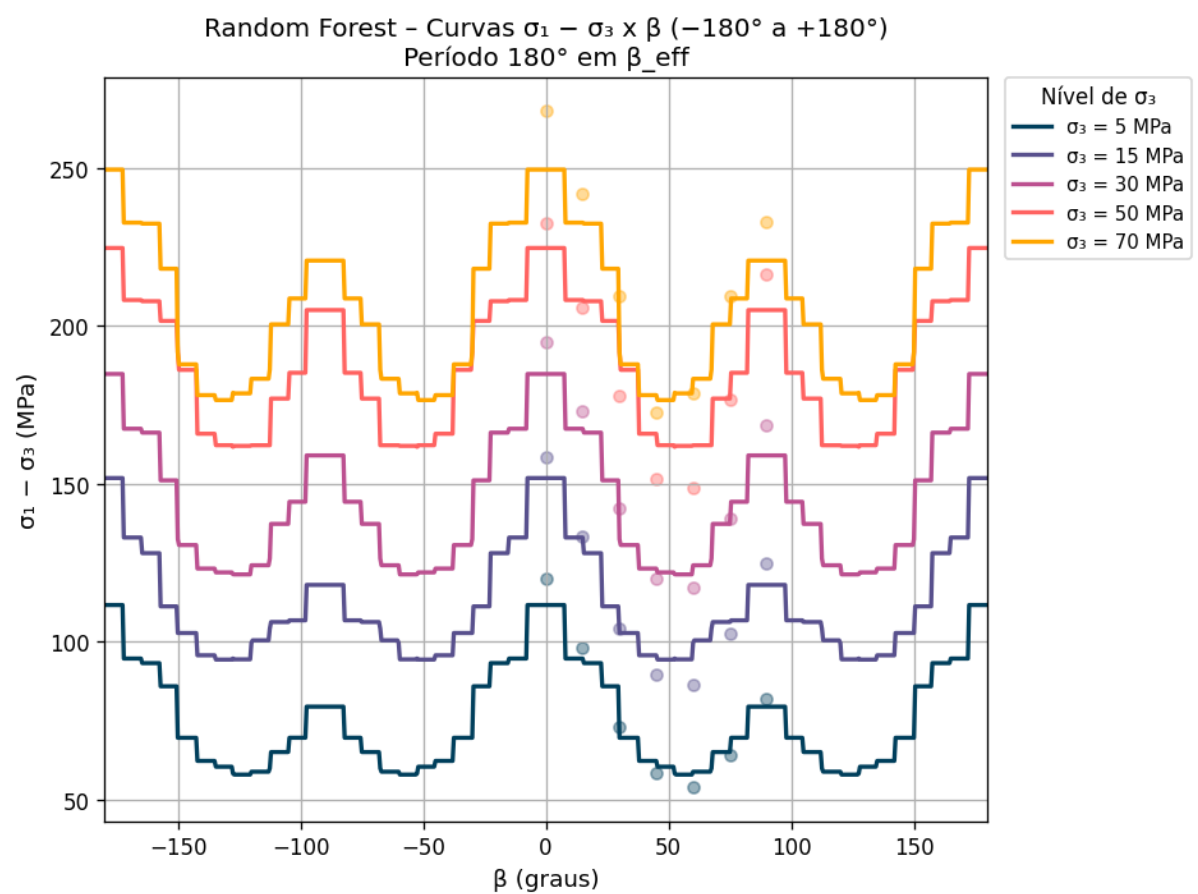
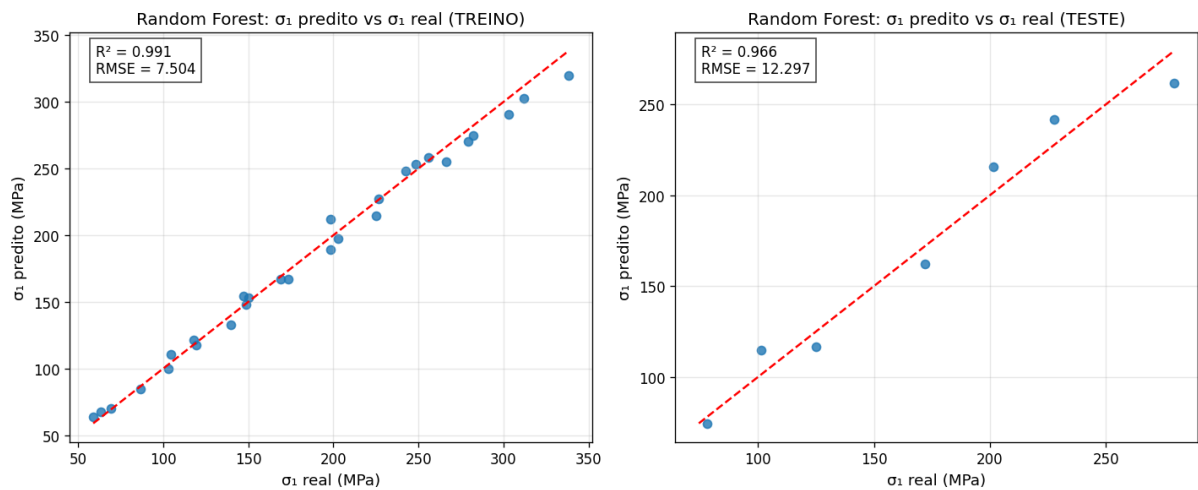
$(\sigma_1 - \sigma_3)$ vs β — Random Forest (0°–360°, período 180°)



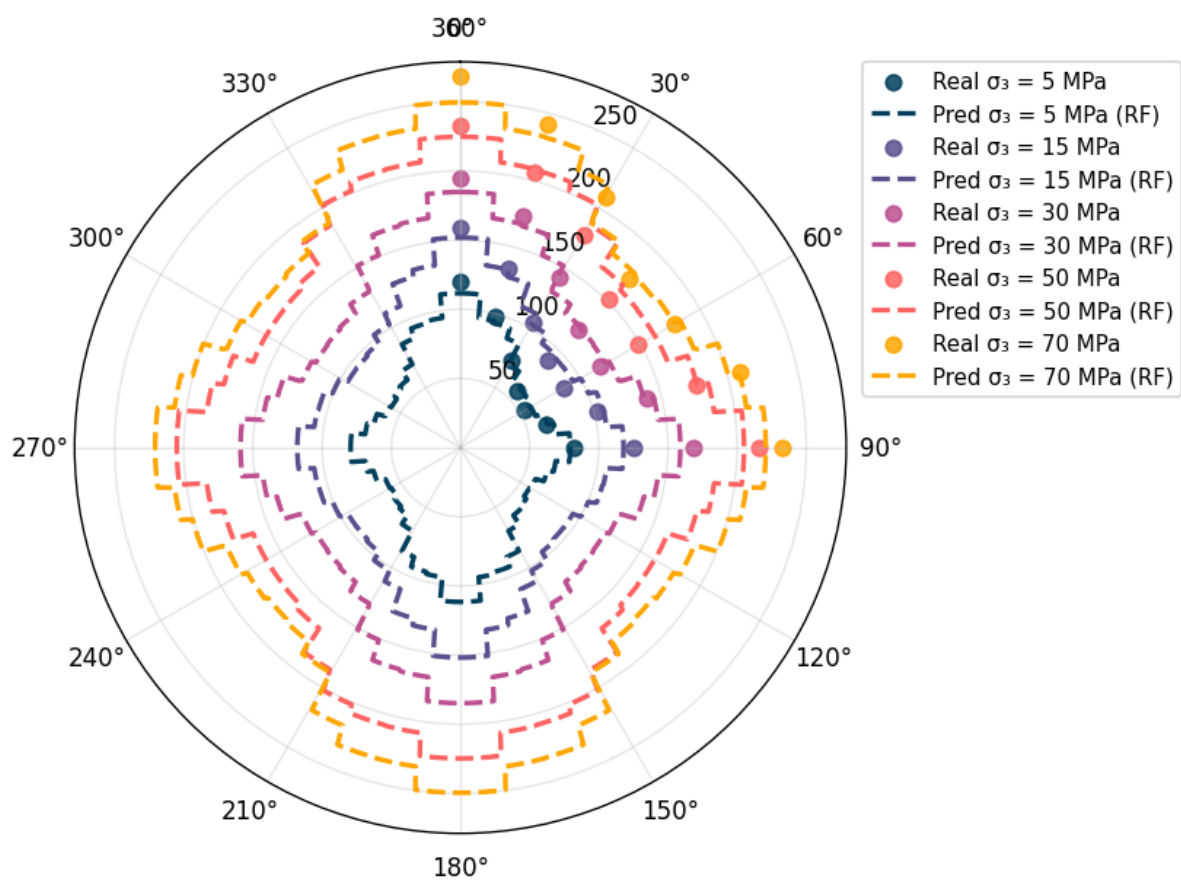
σ_1/σ_3 vs β — Random Forest (0°–360°, período 180°)



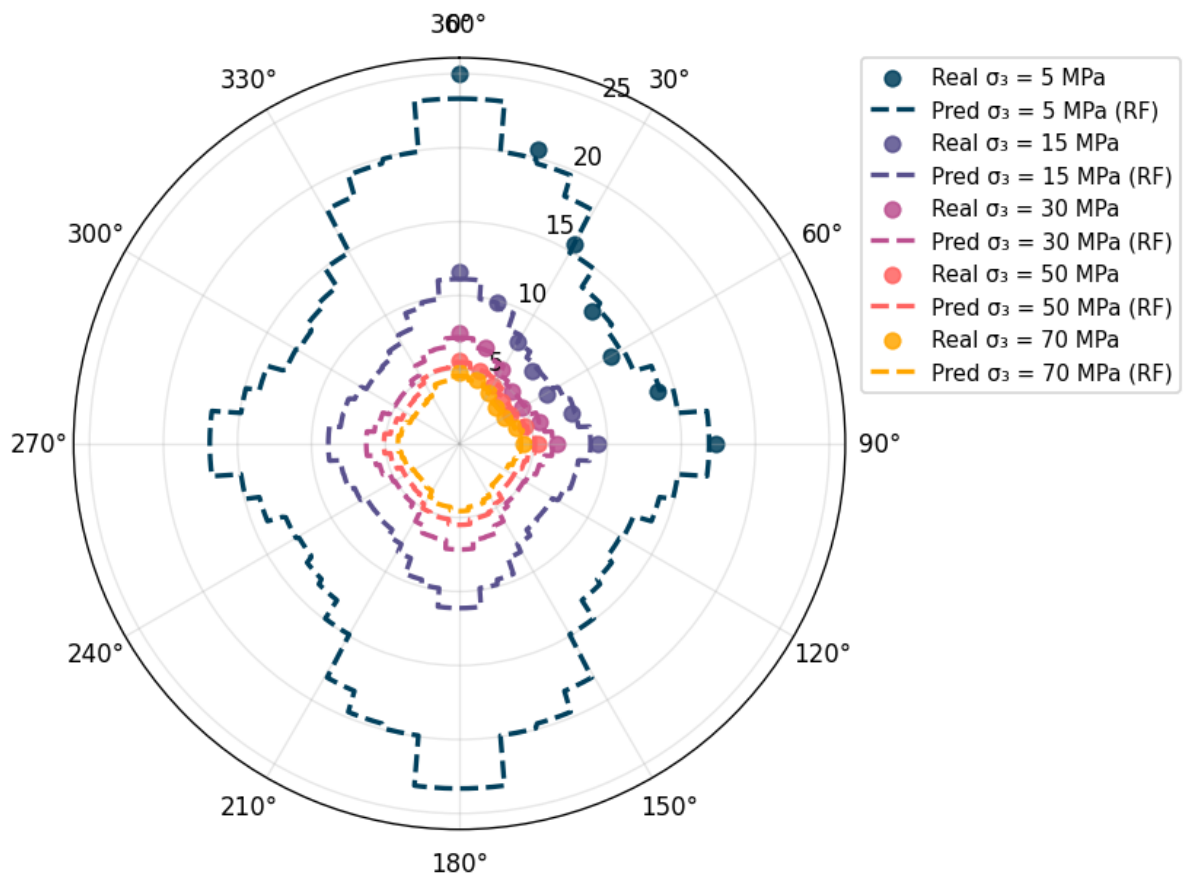
- *Quartzite:*



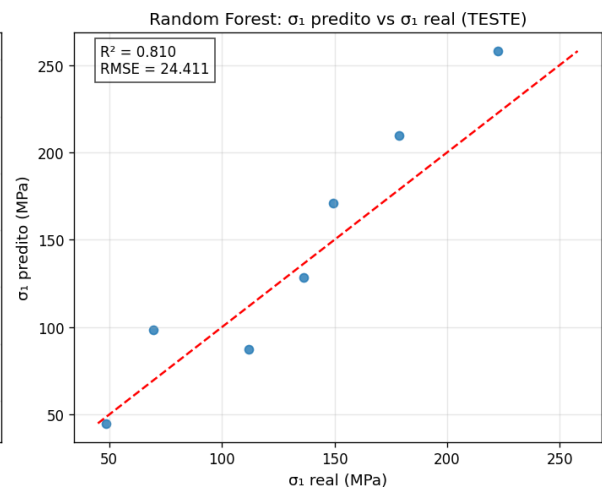
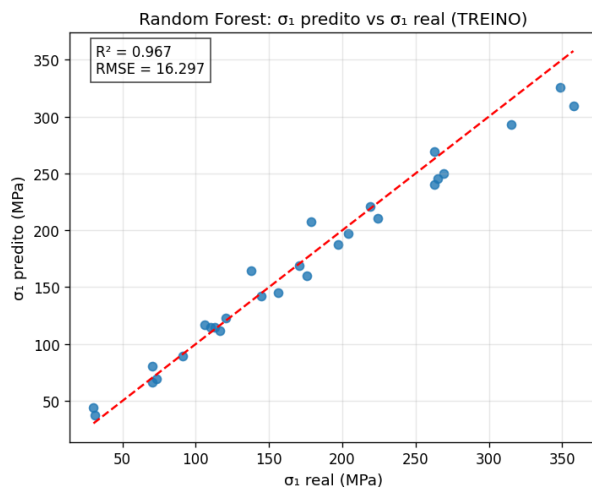
$(\sigma_1 - \sigma_3)$ vs β — Random Forest (0°–360°, período 180°)

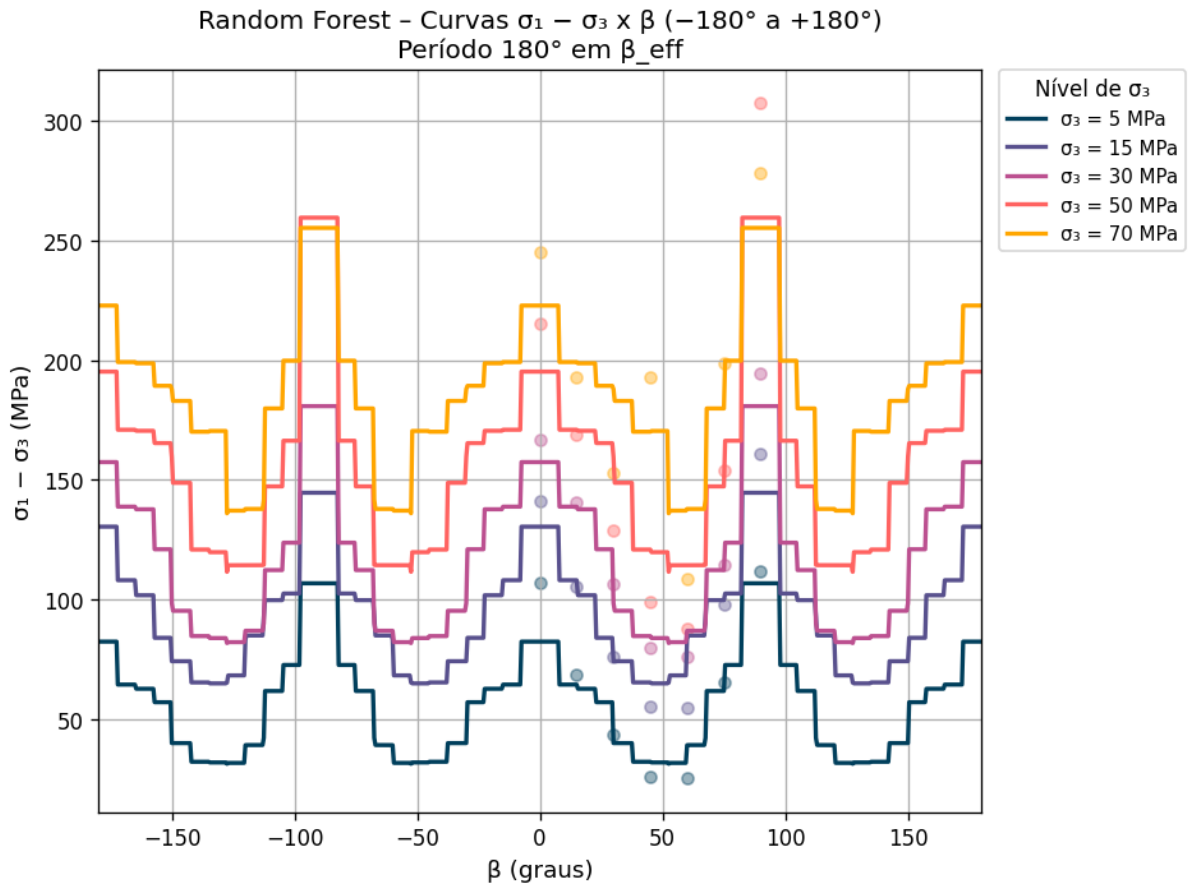


σ_1/σ_3 vs β — Random Forest (0°–360°, período 180°)

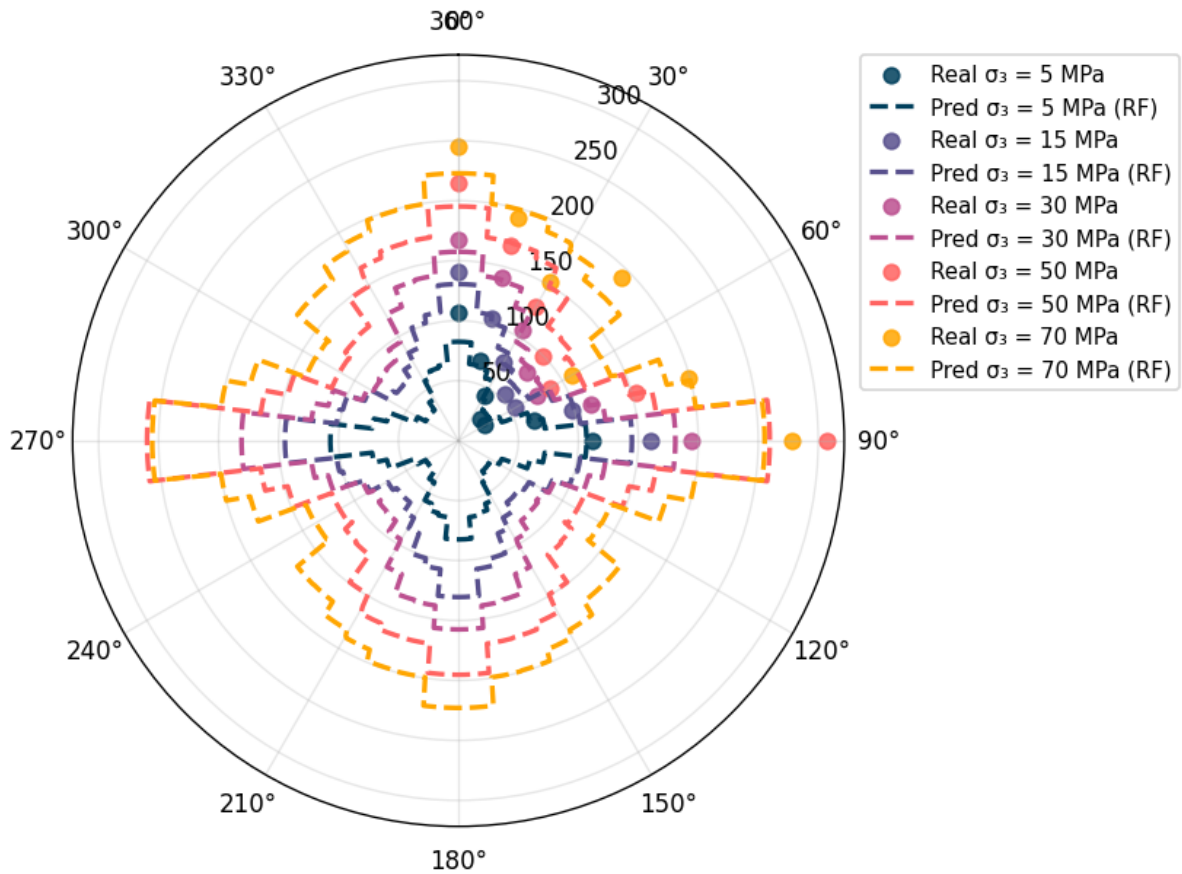


• *Micaceous:*

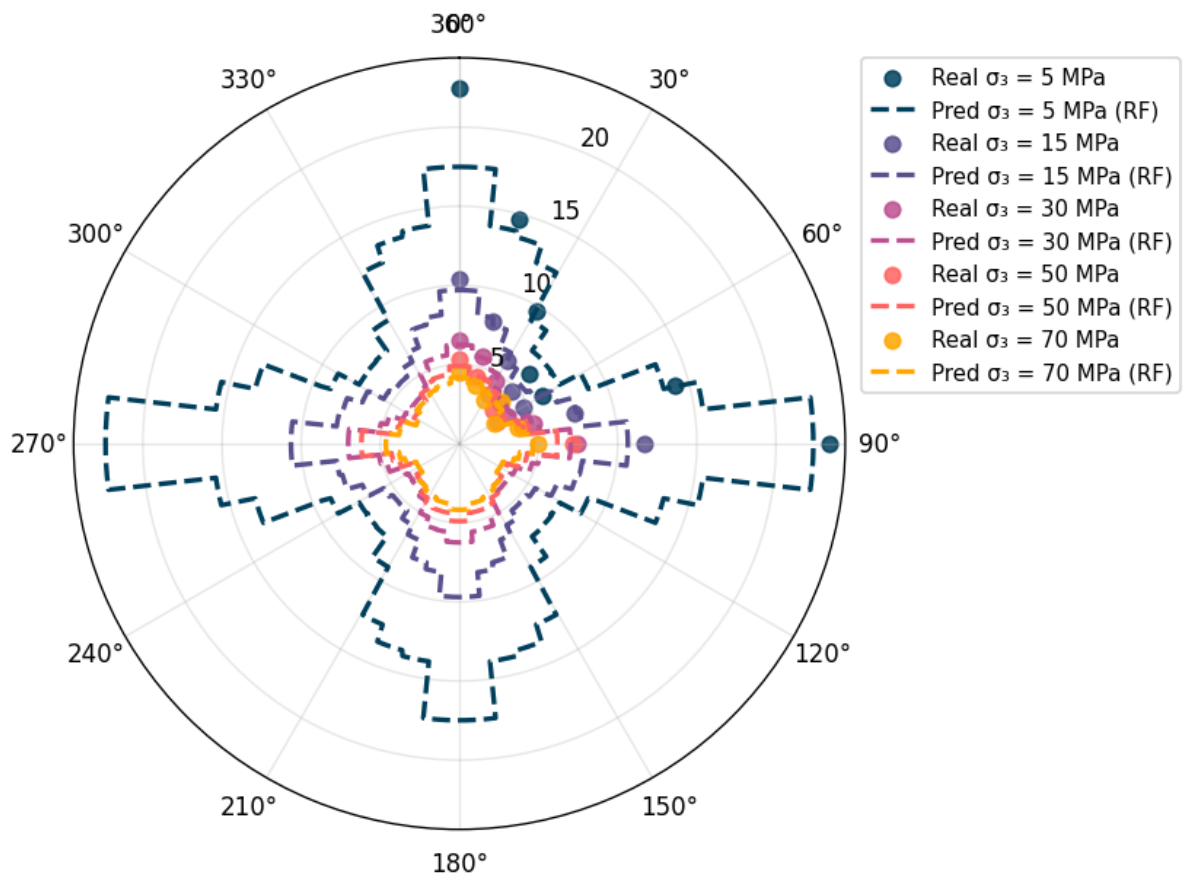




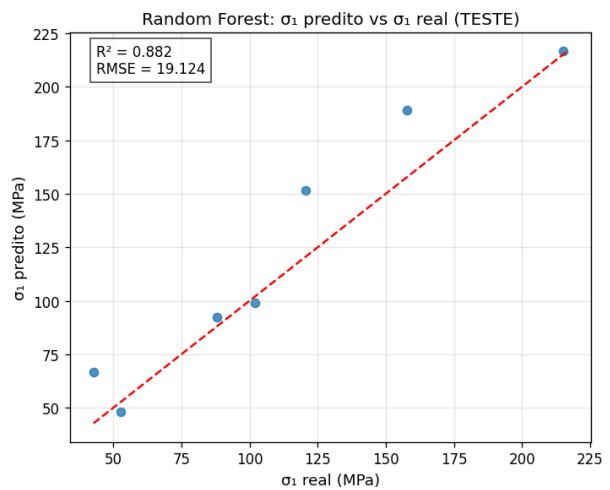
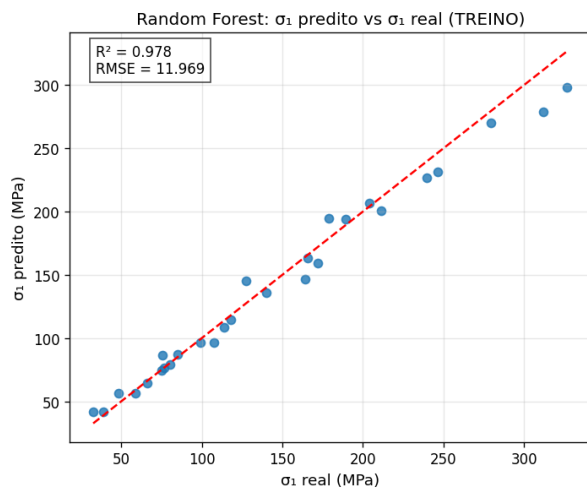
$(\sigma_1 - \sigma_3)$ vs β — Random Forest (0° - 360° , período 180°)

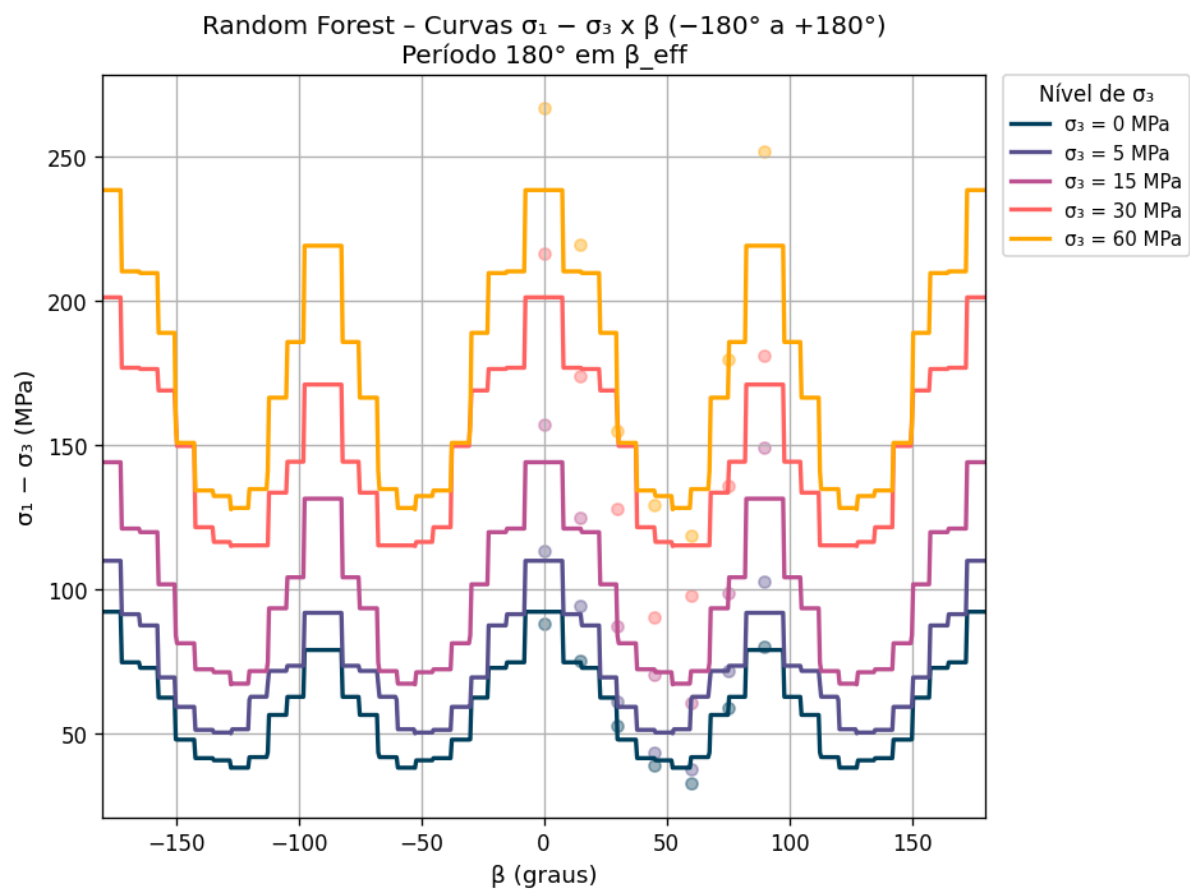


σ_1/σ_3 vs β — Random Forest (0°–360°, período 180°)

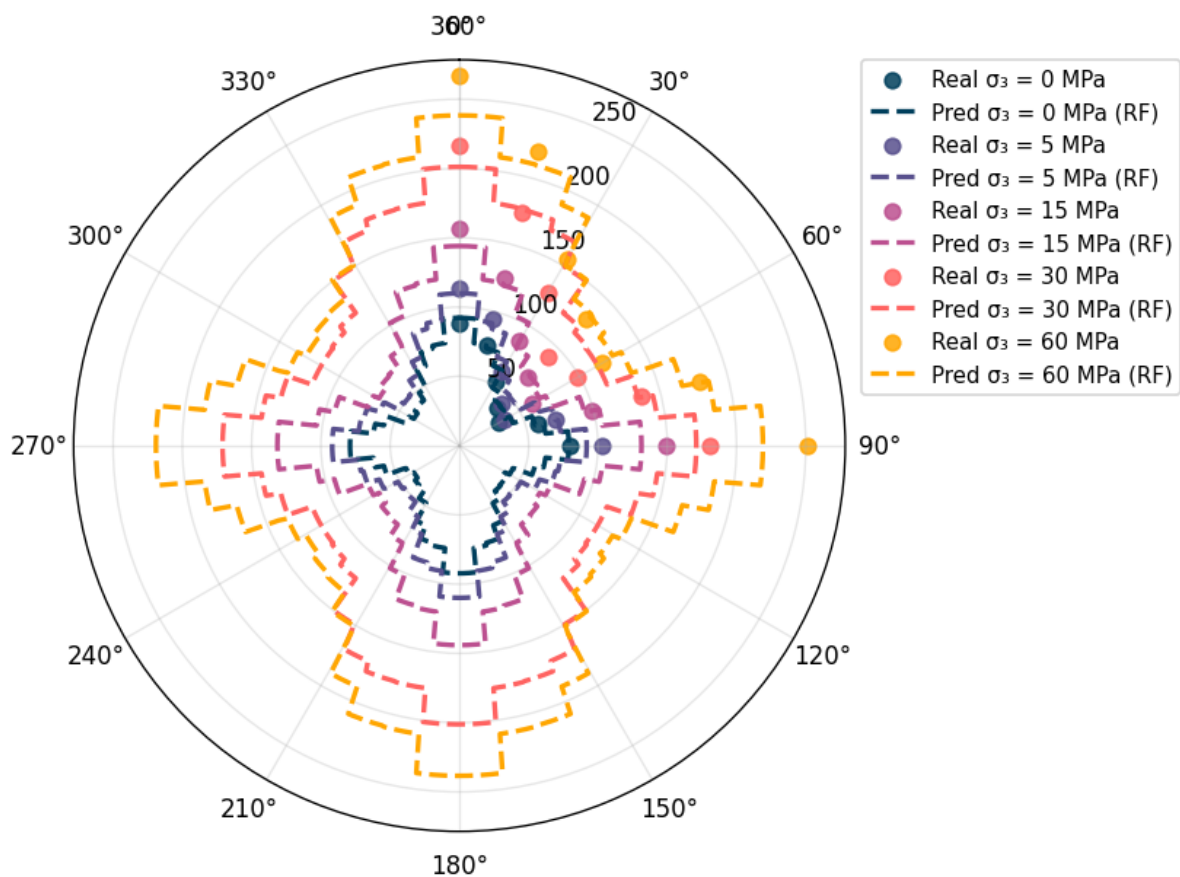


• *Dry:*

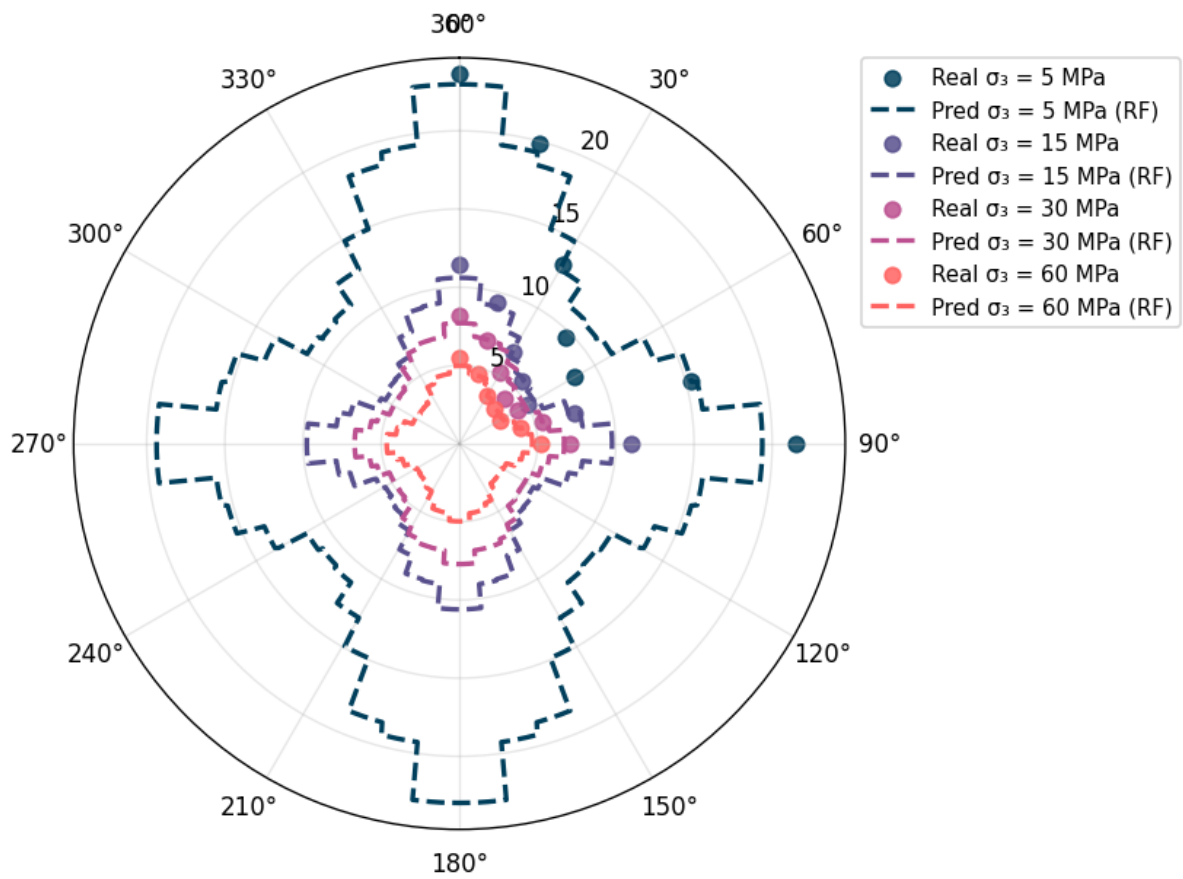




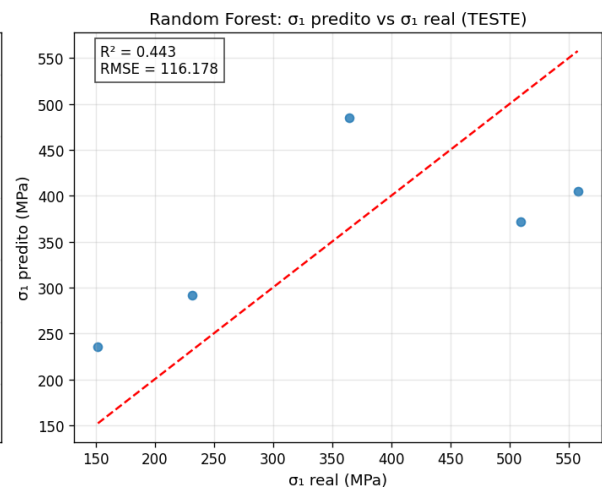
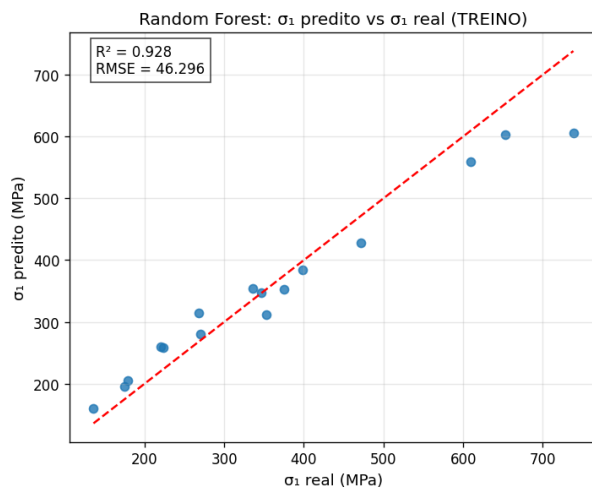
$(\sigma_1 - \sigma_3)$ vs β — Random Forest (0° - 360° , período 180°)

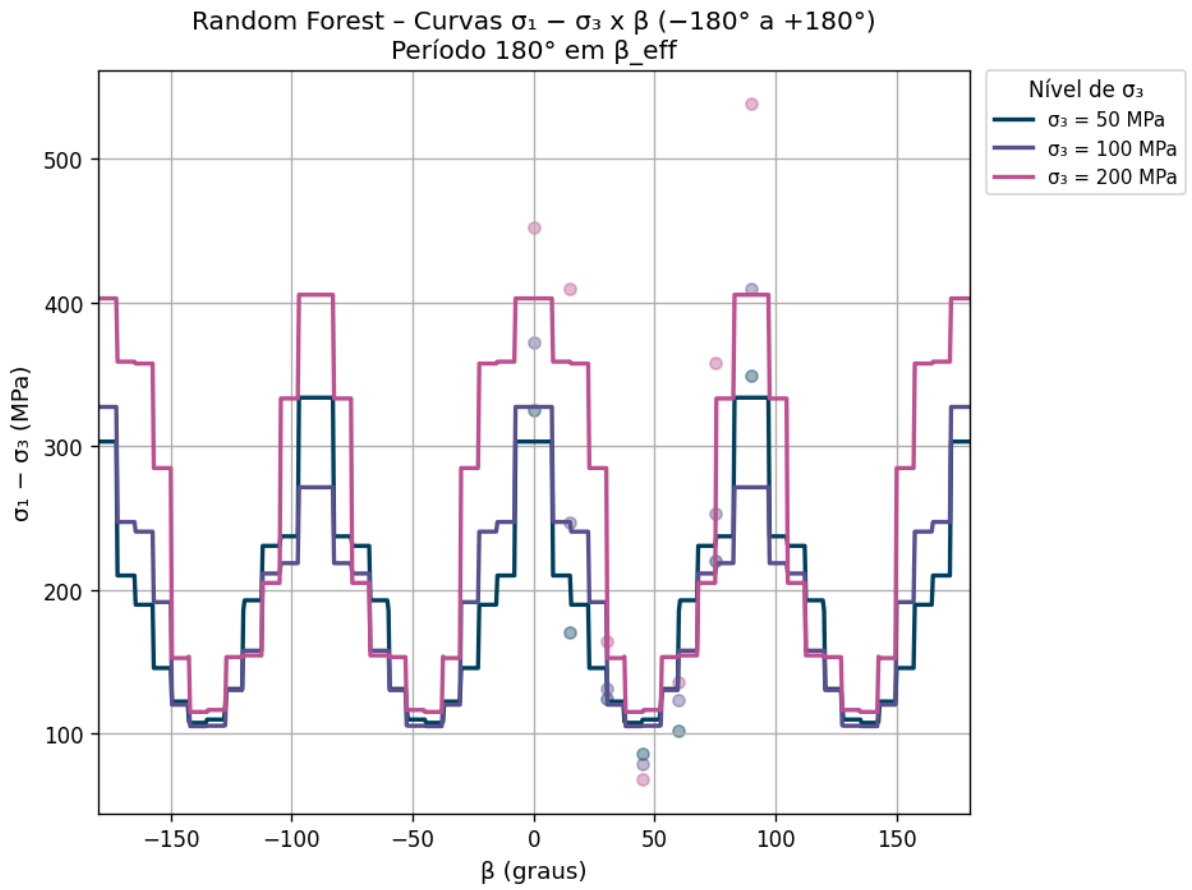


σ_1/σ_3 vs β — Random Forest (0°–360°, período 180°)

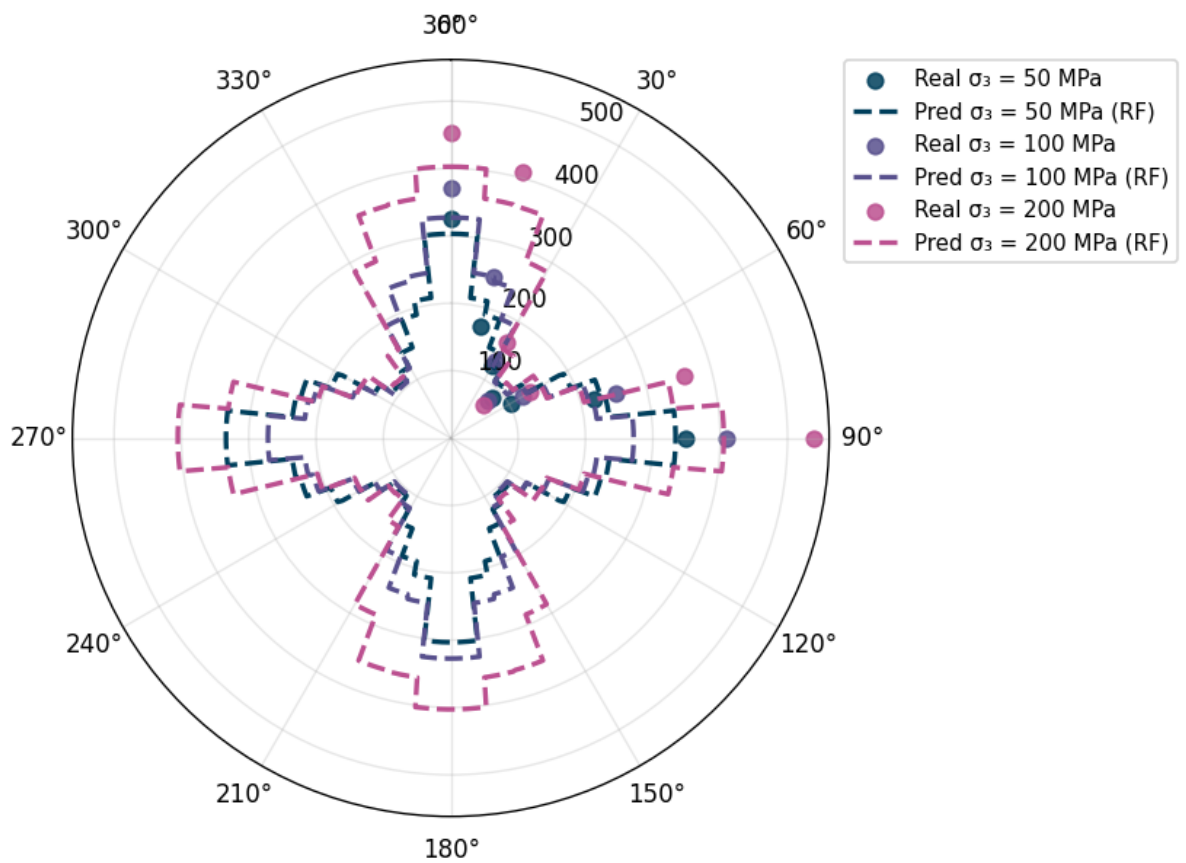


• *Moretown:*

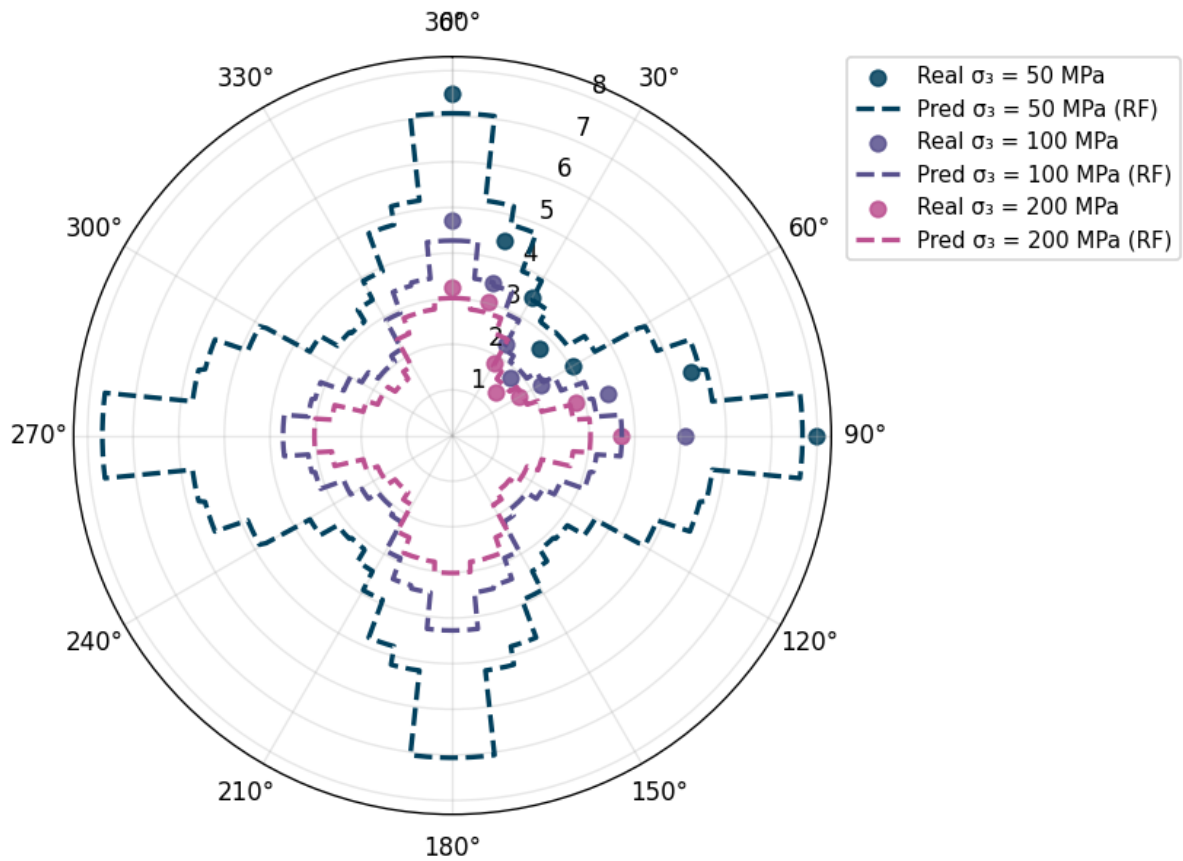




$(\sigma_1 - \sigma_3)$ vs β — Random Forest (0° – 360° , período 180°)

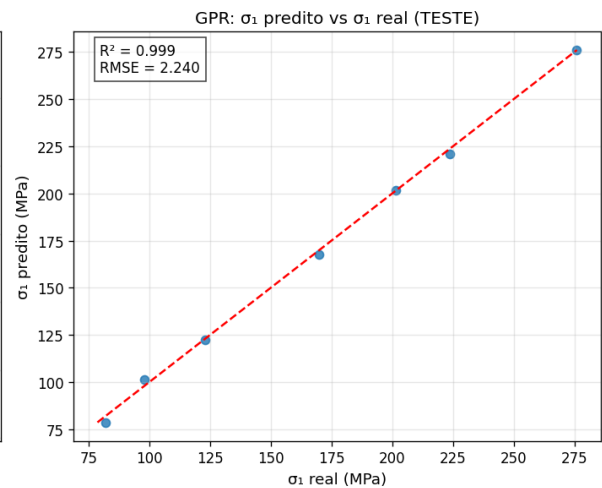
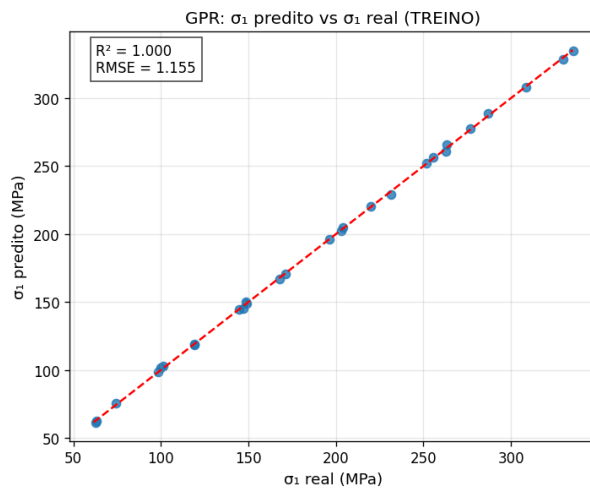


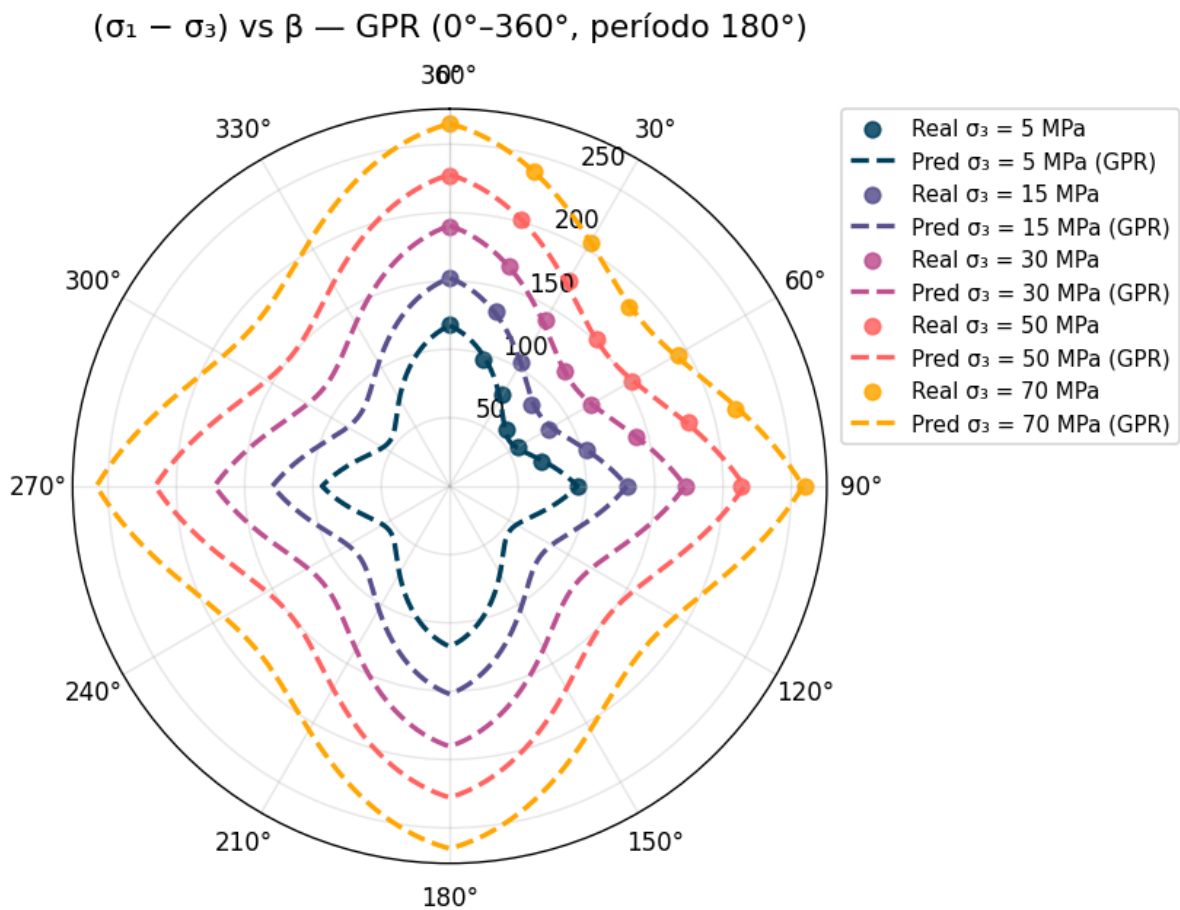
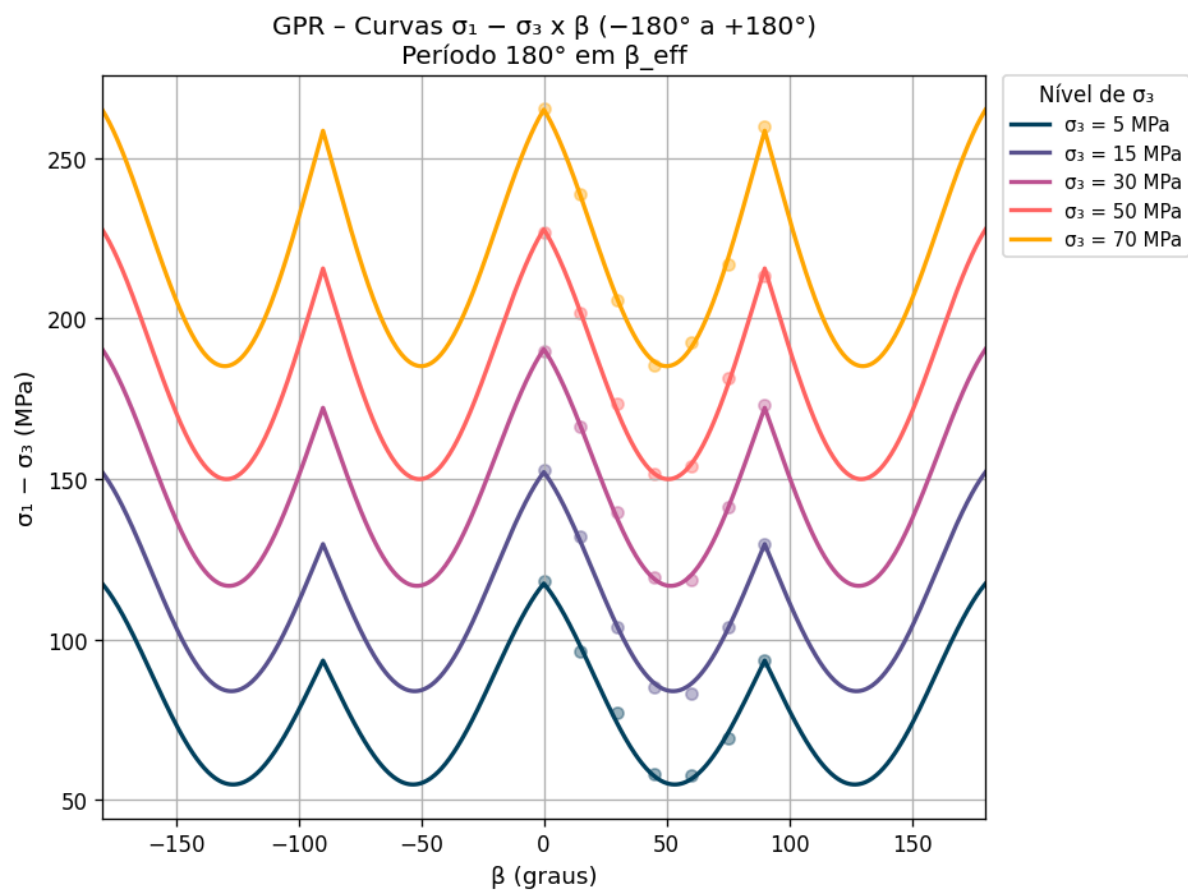
σ_1/σ_3 vs β — Random Forest (0°-360°, período 180°)

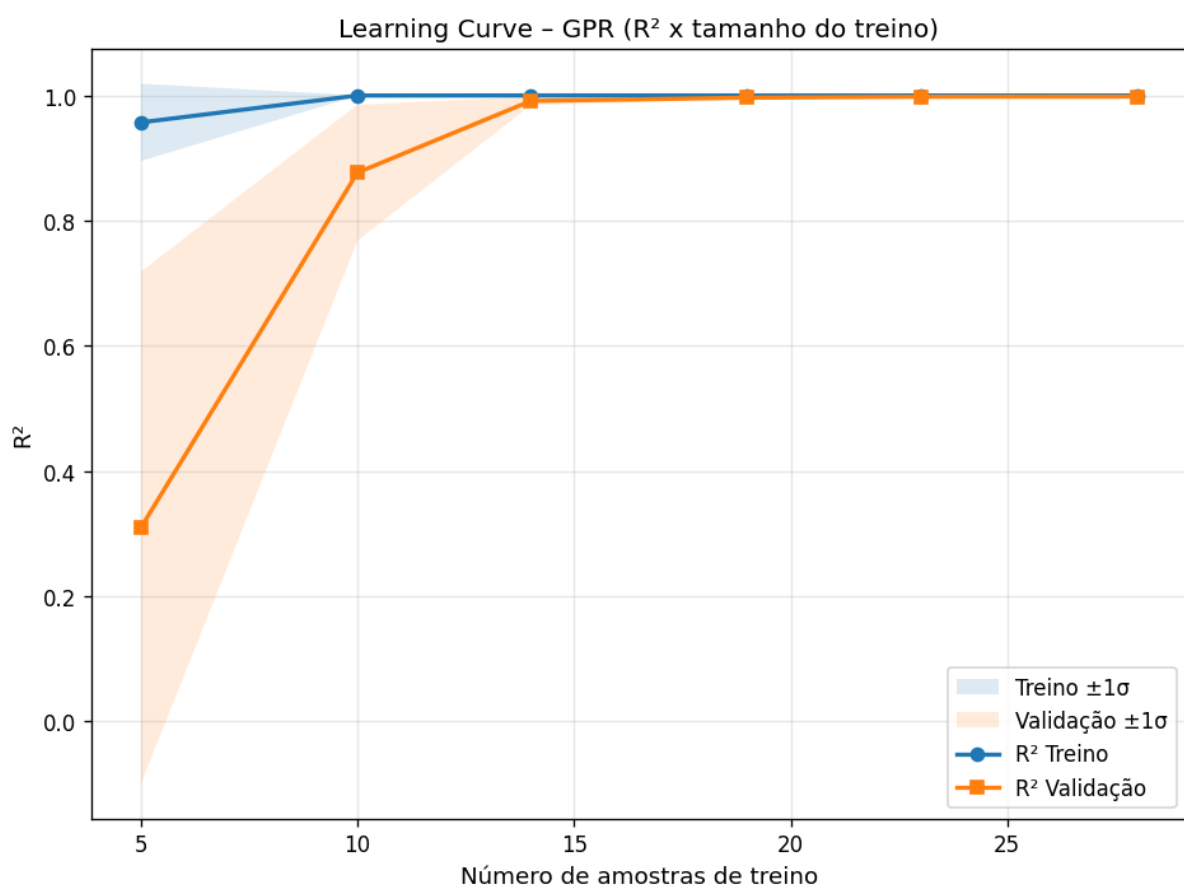
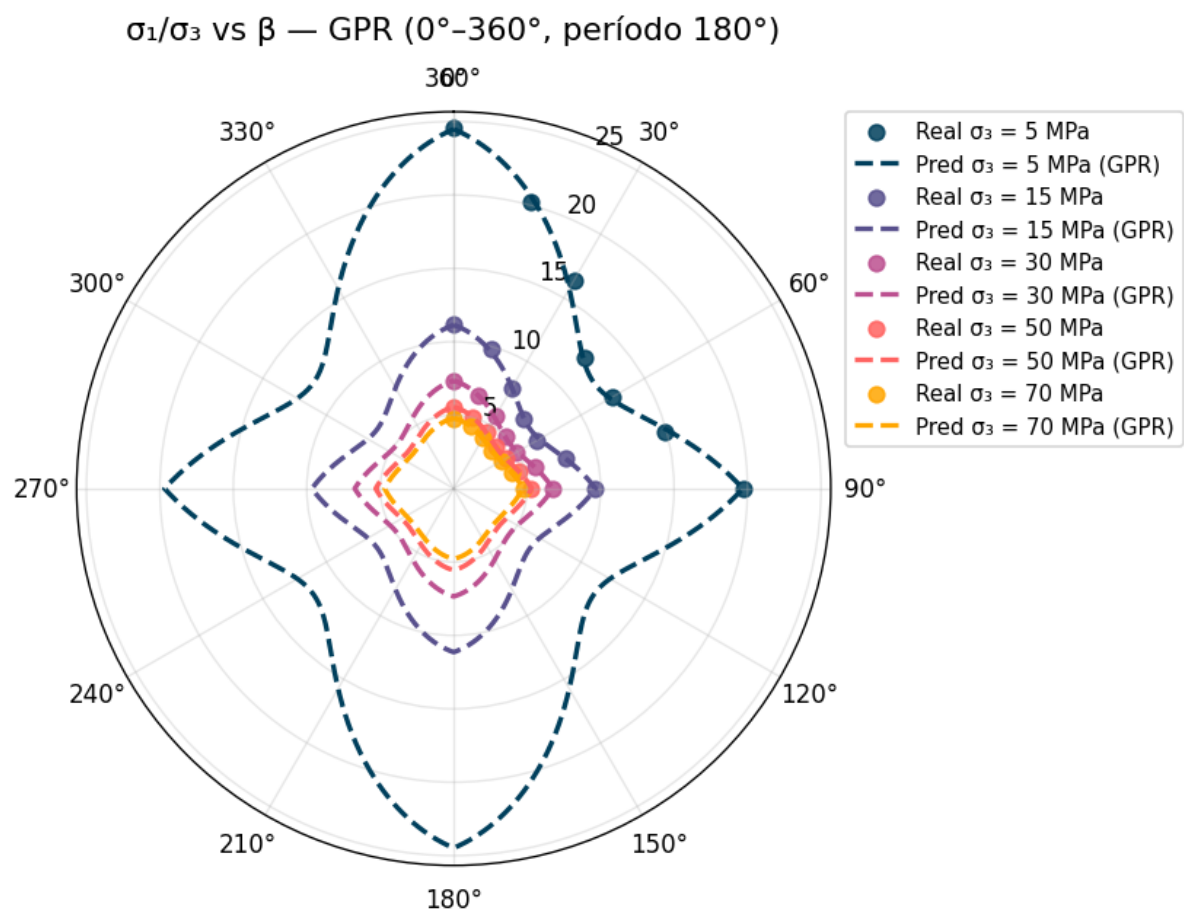


2.3. Resultados do GPR:

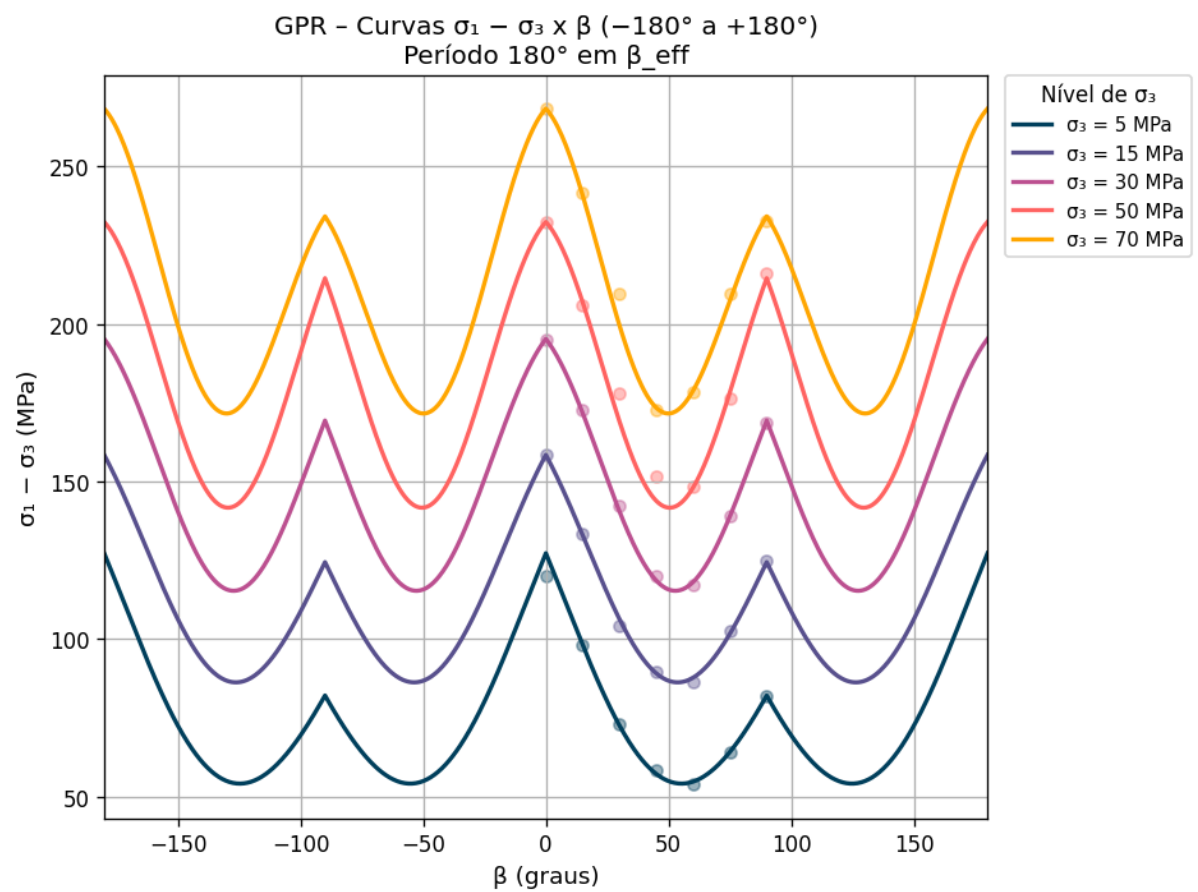
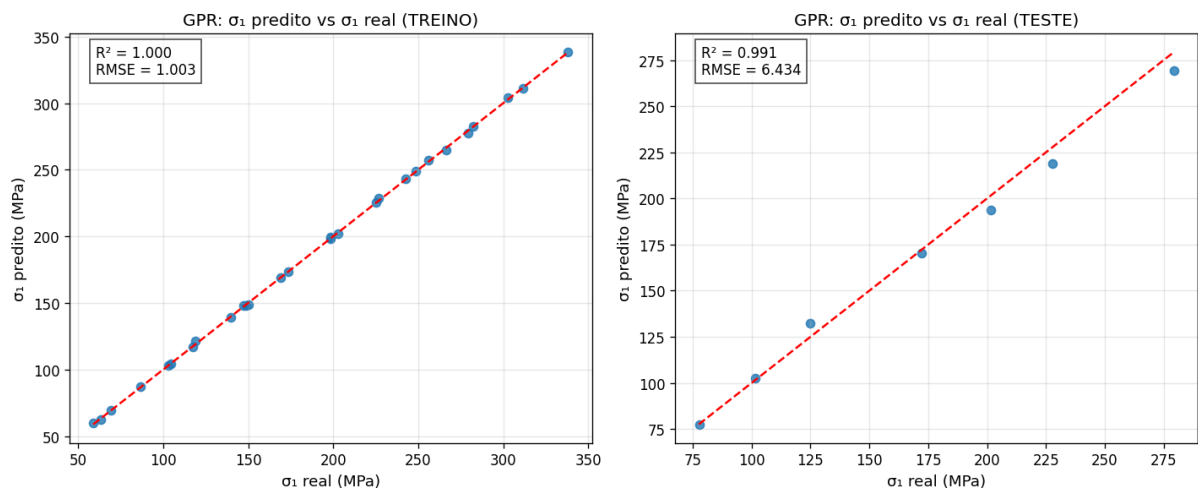
- *Carbonaceous:*



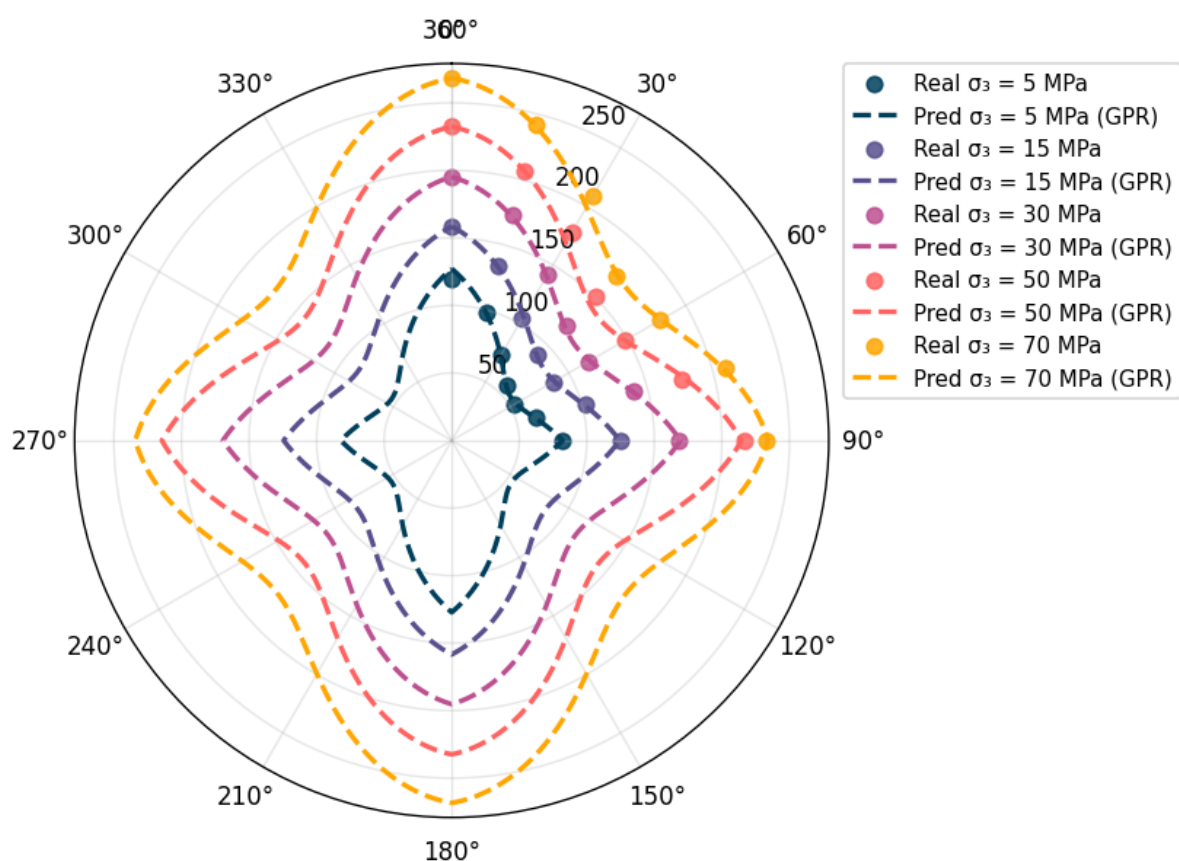




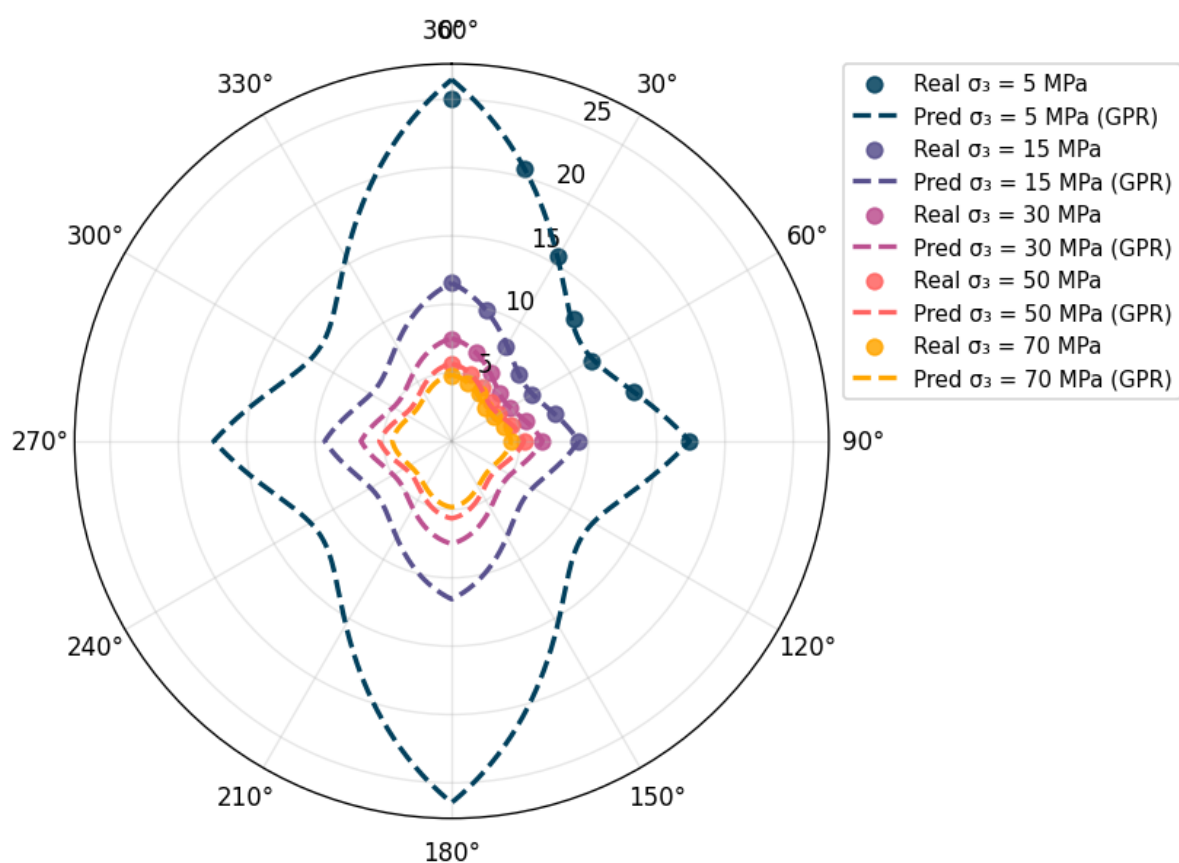
- *Quartzite:*

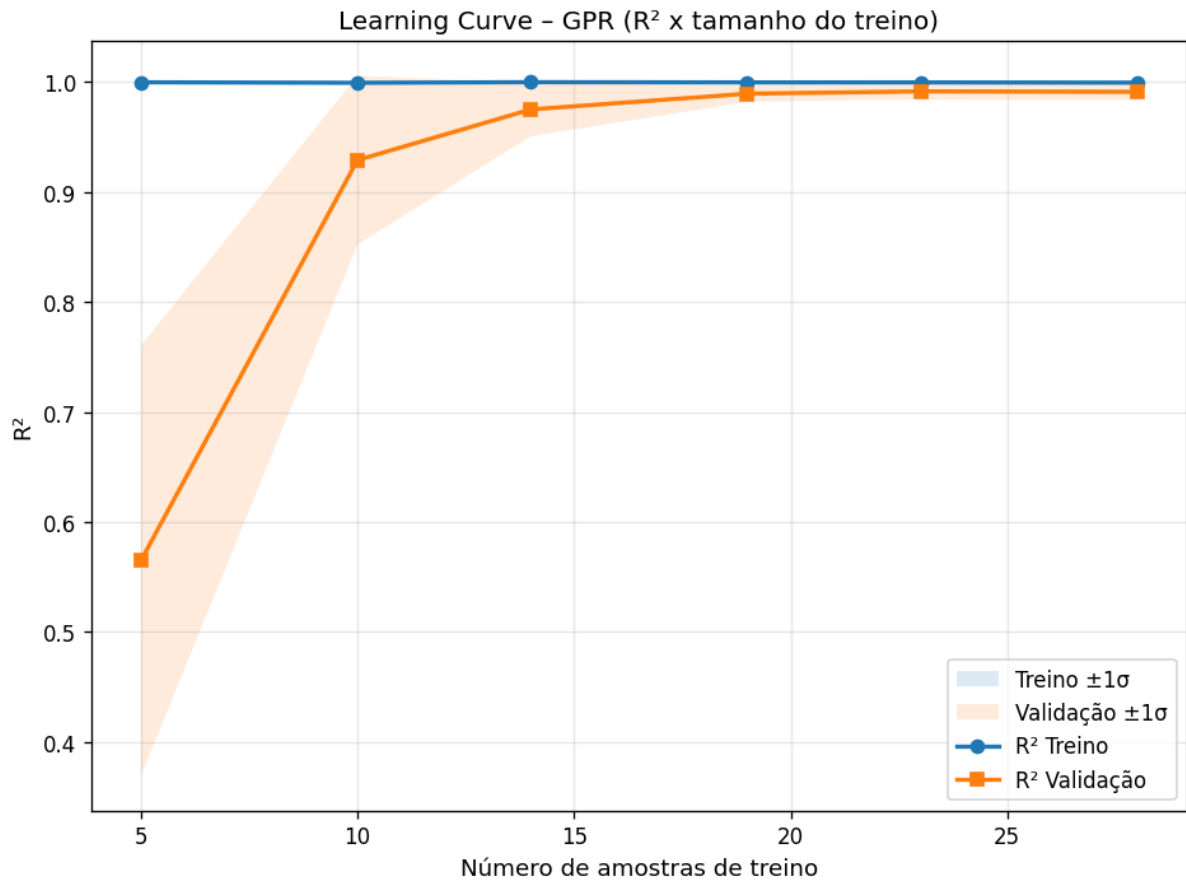


$(\sigma_1 - \sigma_3)$ vs β — GPR (0°–360°, período 180°)

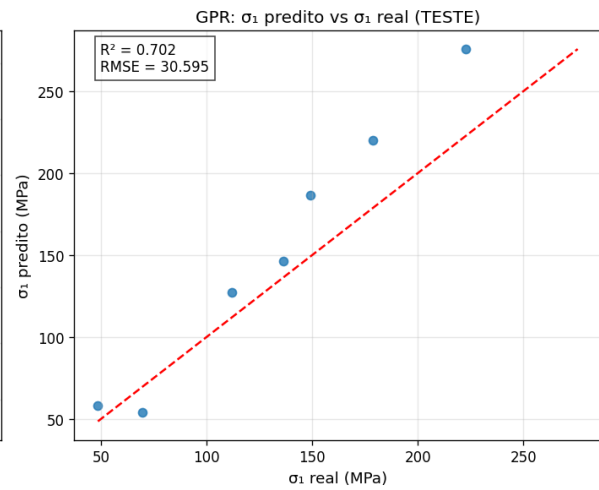
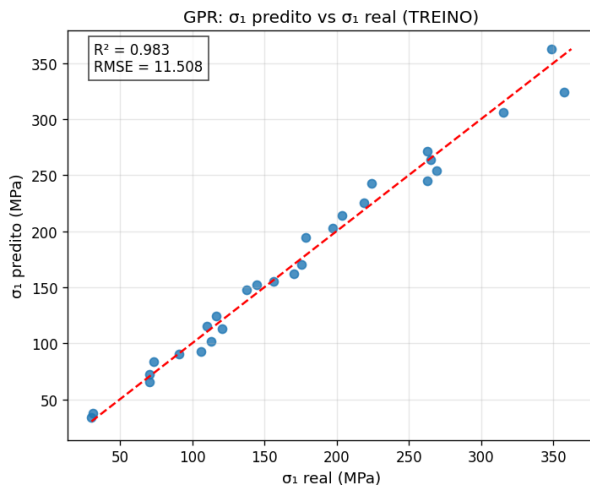


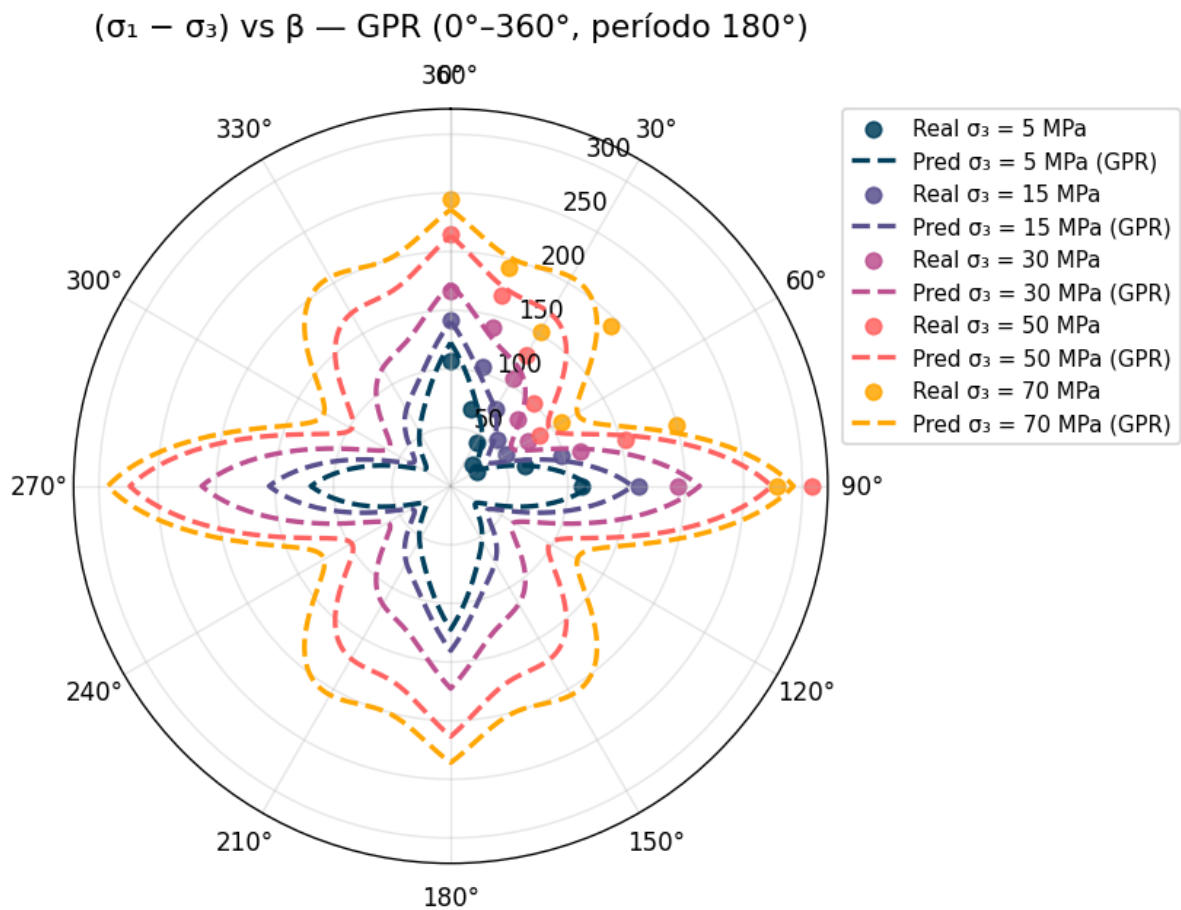
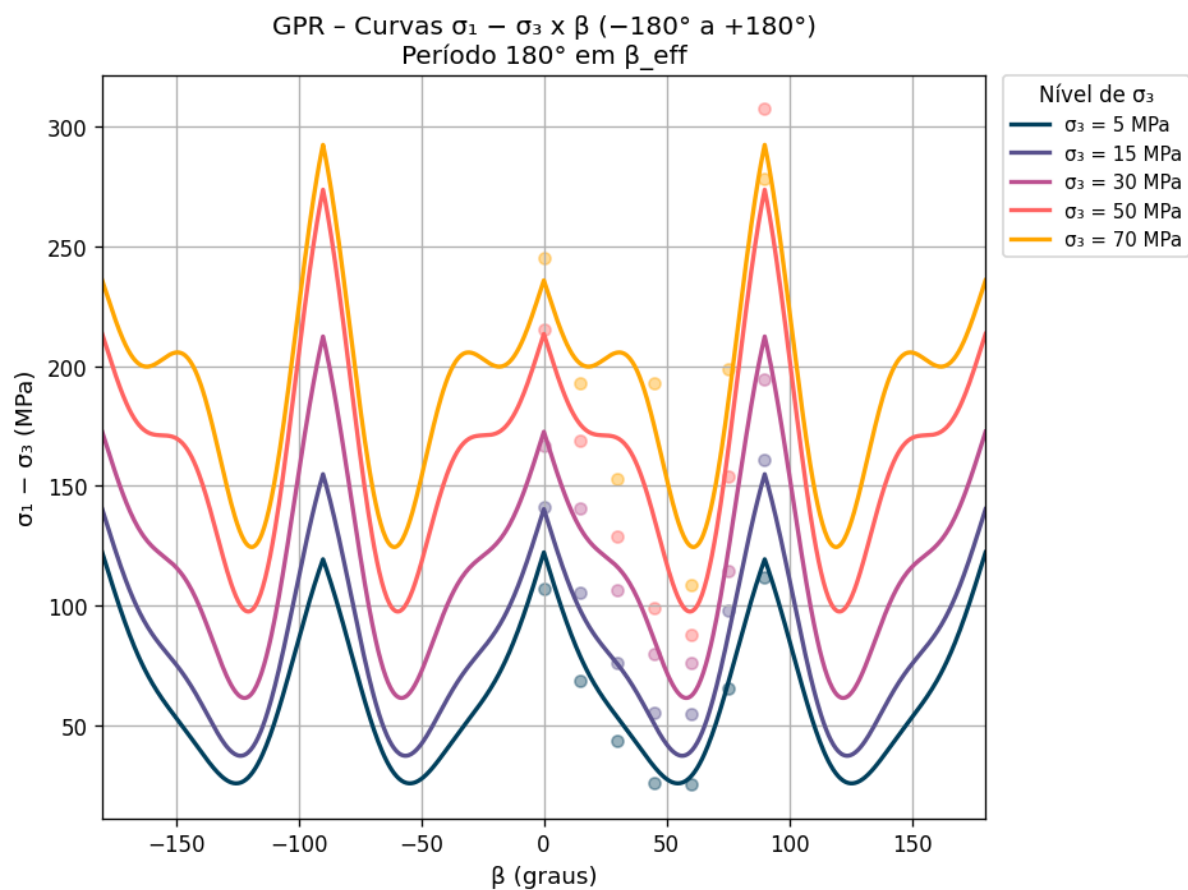
σ_1/σ_3 vs β — GPR (0°–360°, período 180°)

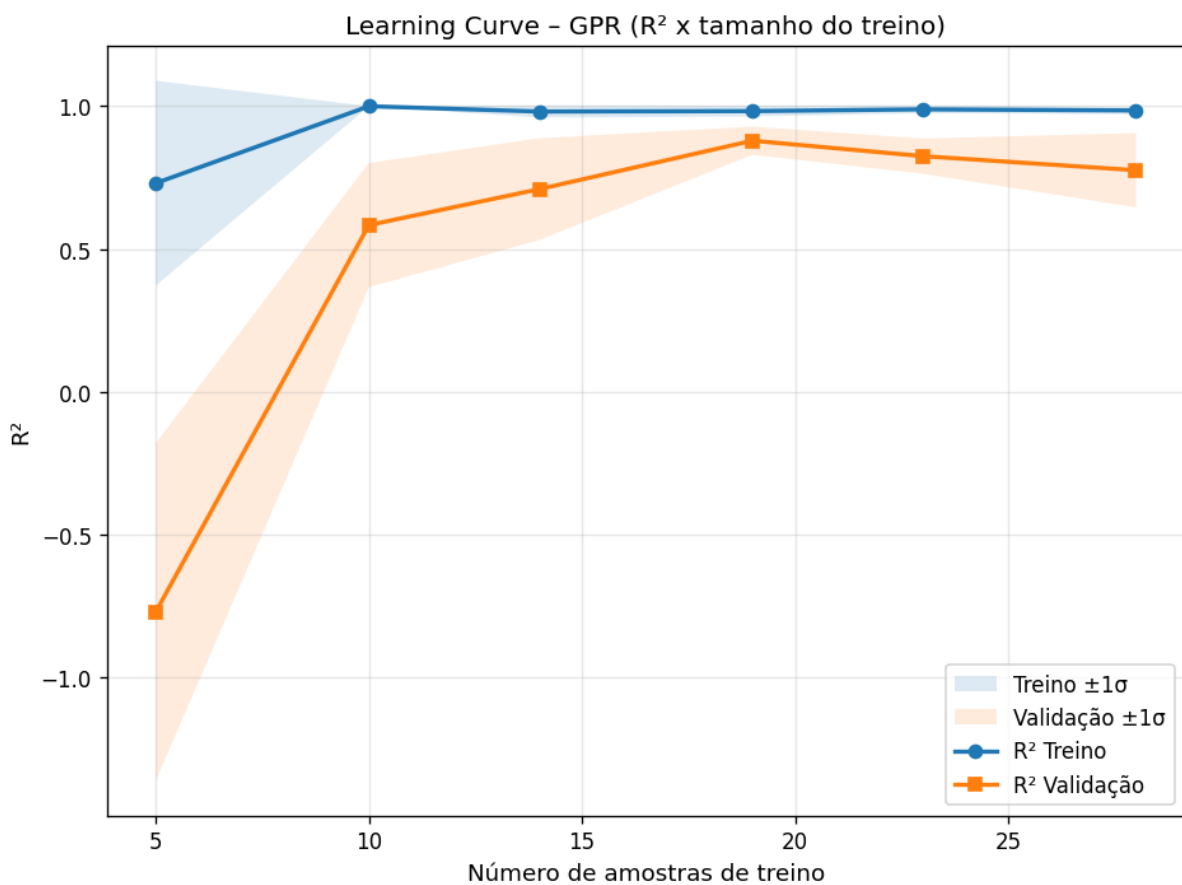
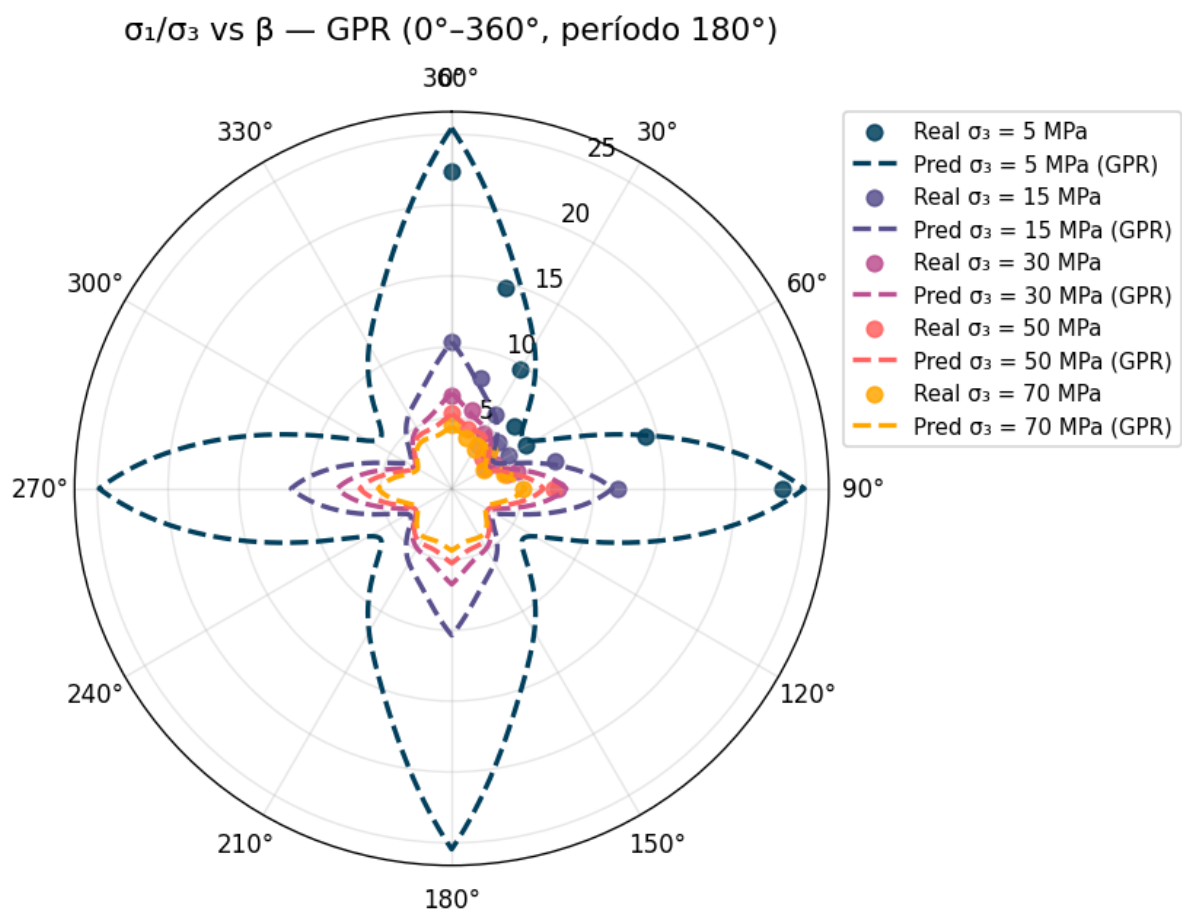




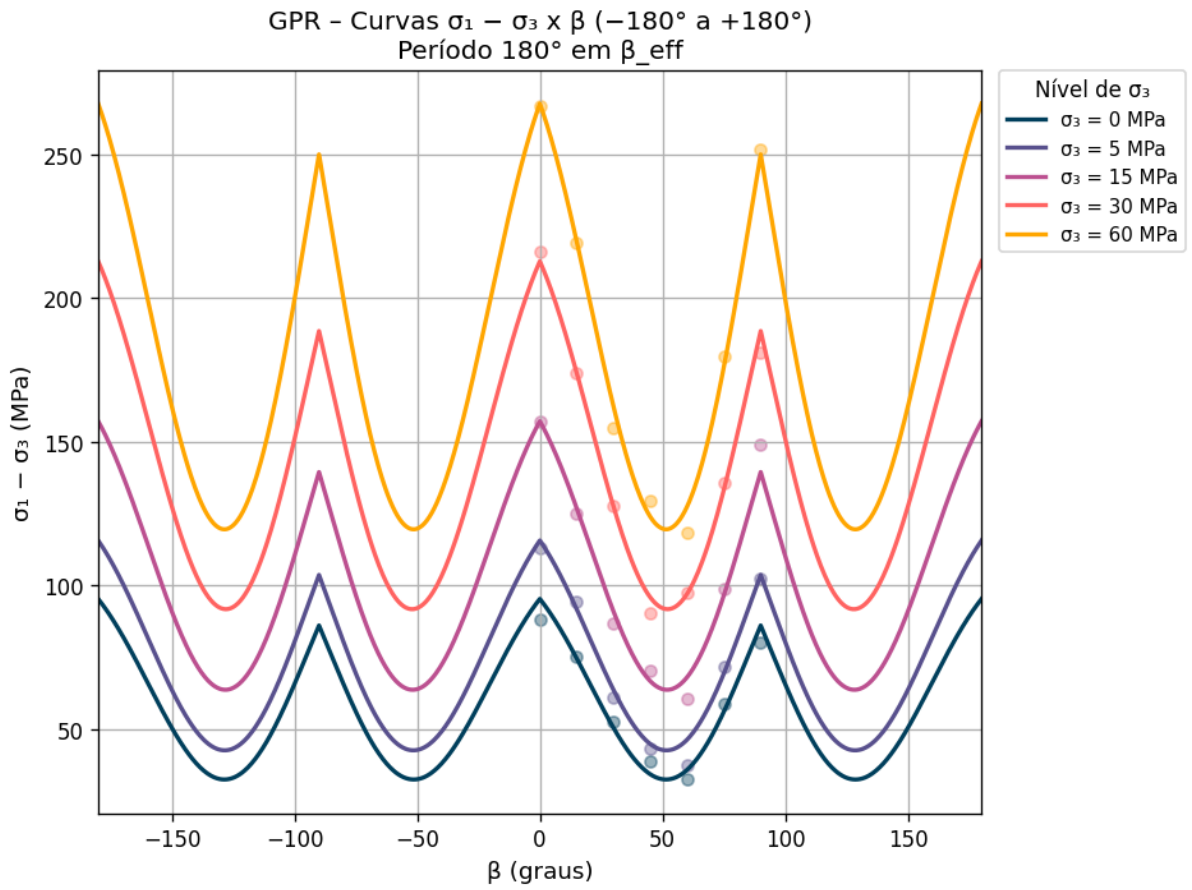
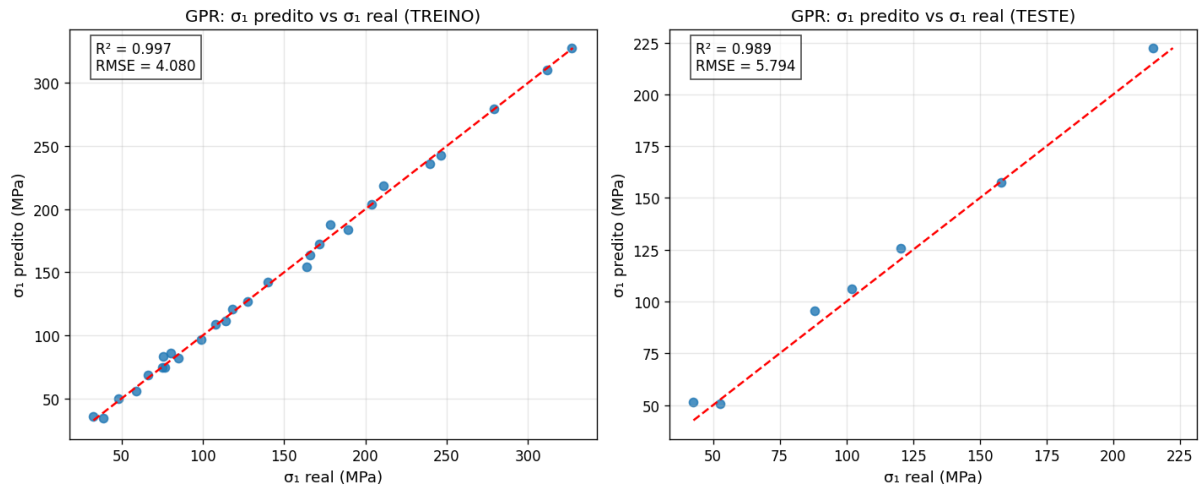
- Micaceous:*



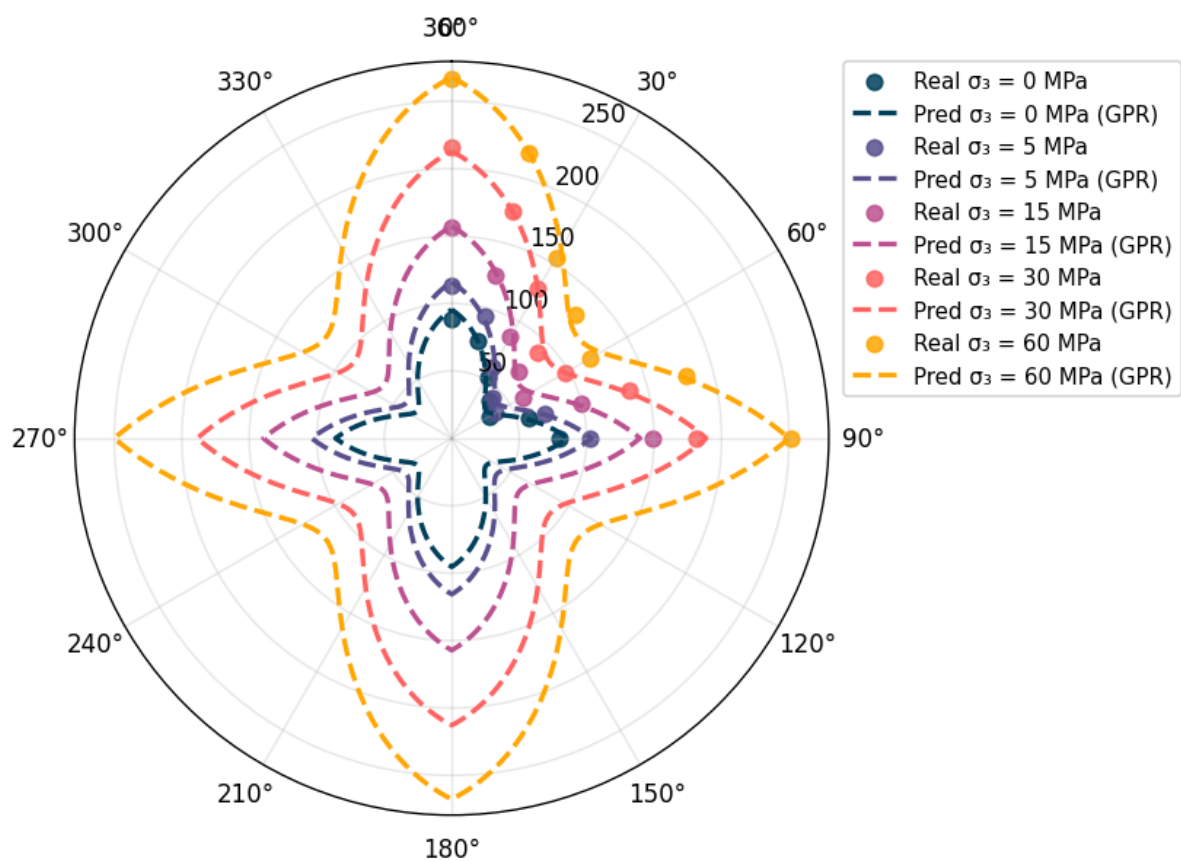




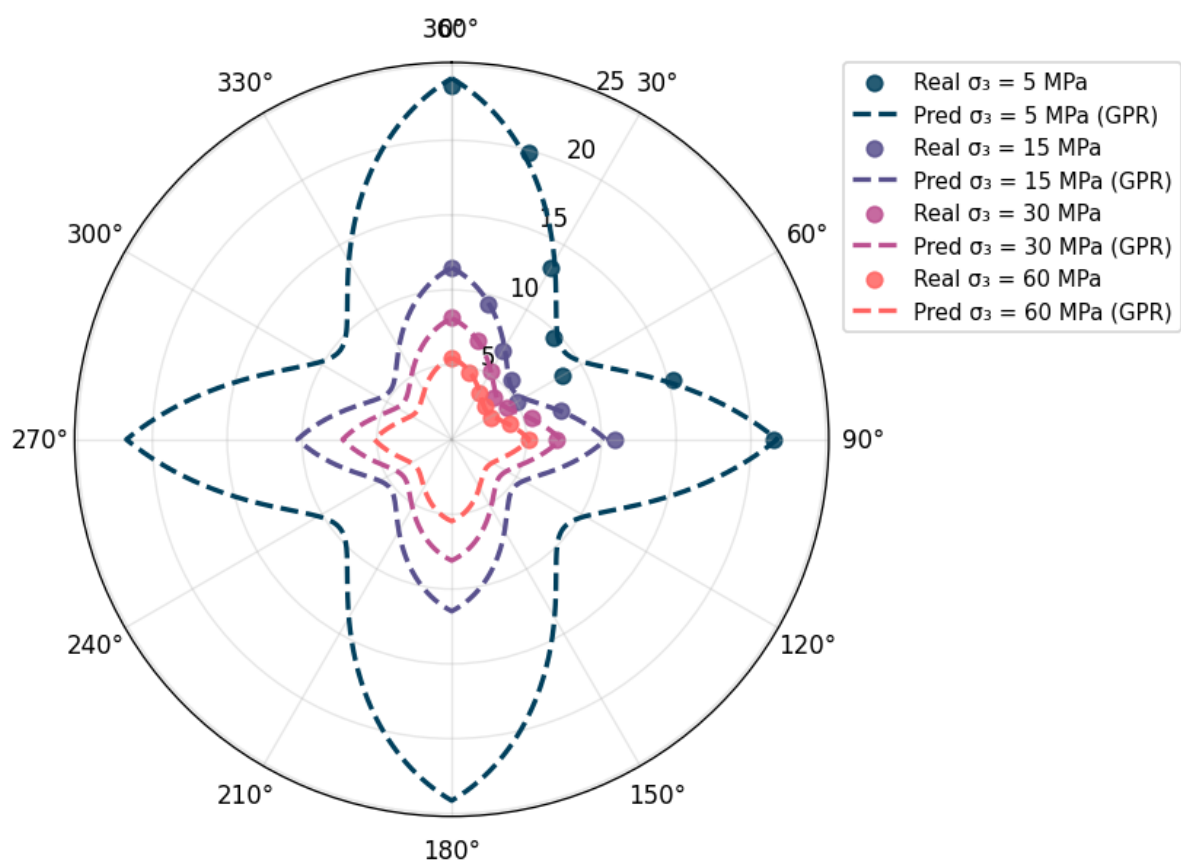
- *Dry:*

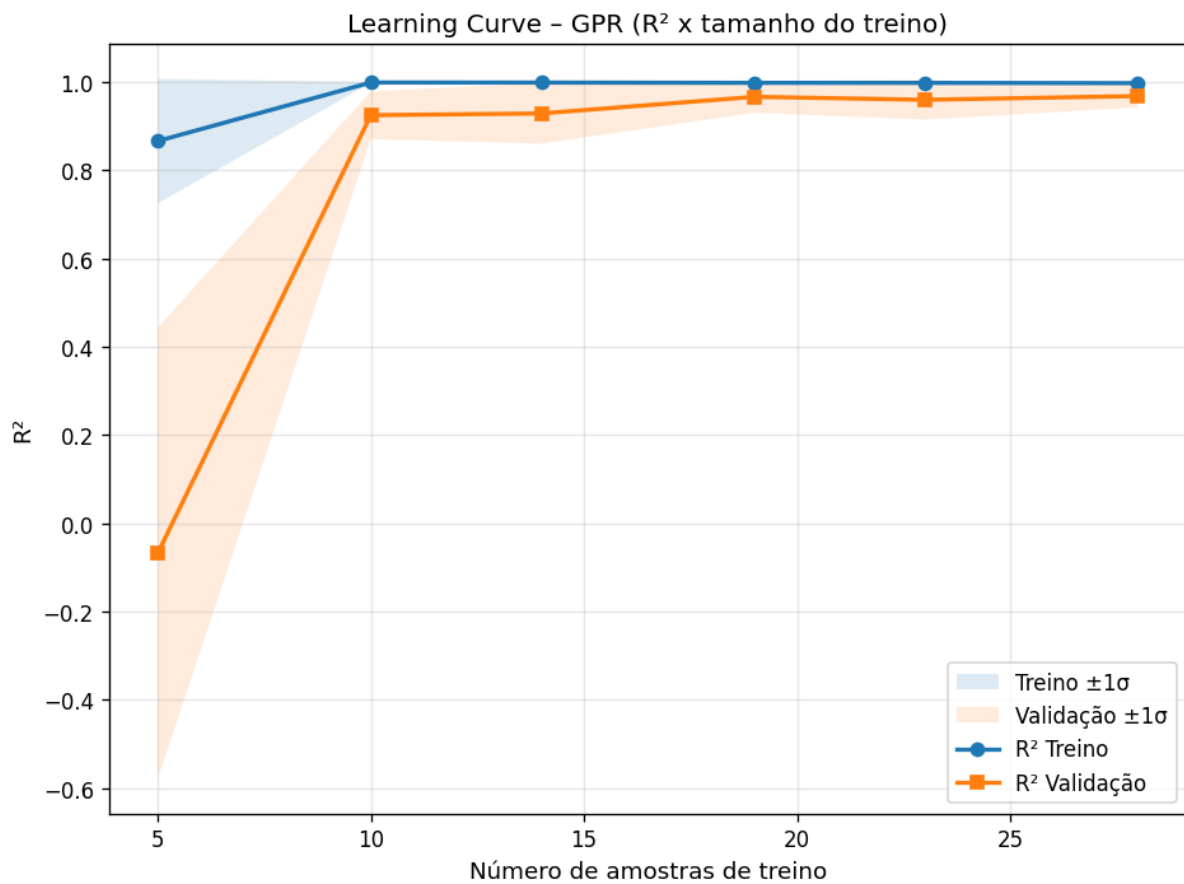


$(\sigma_1 - \sigma_3)$ vs β — GPR (0° – 360° , período 180°)

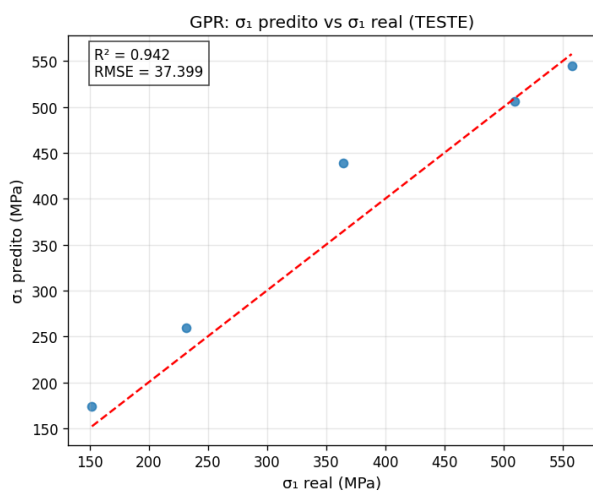
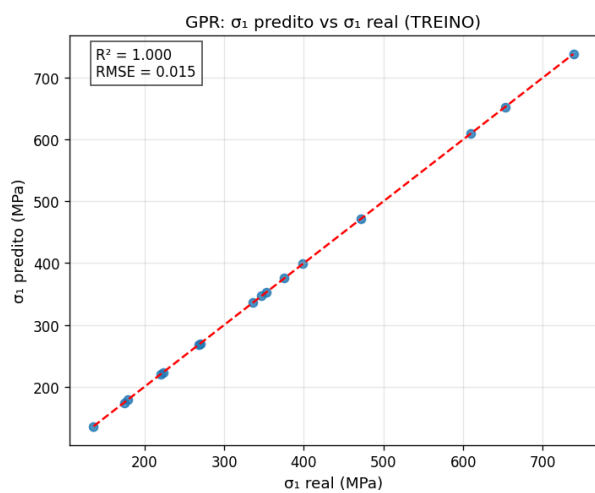


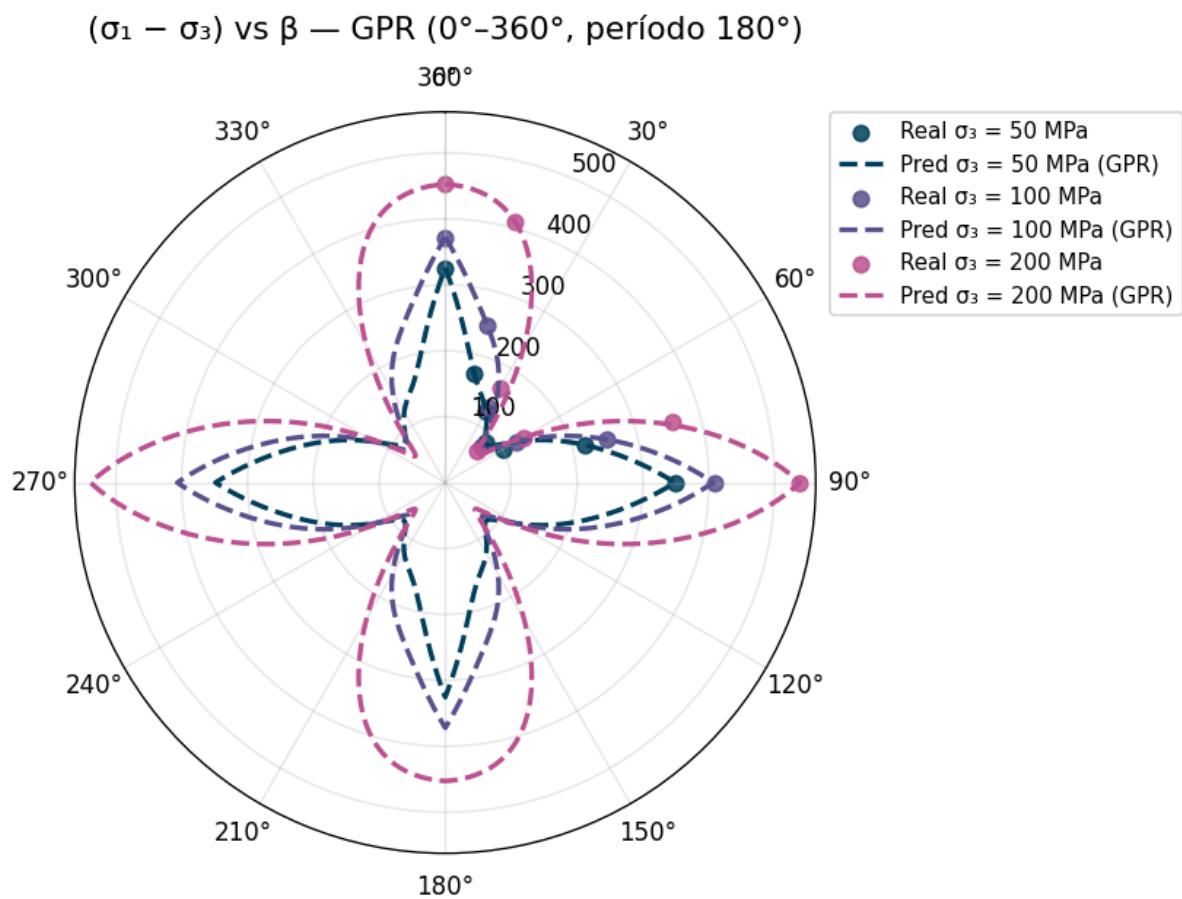
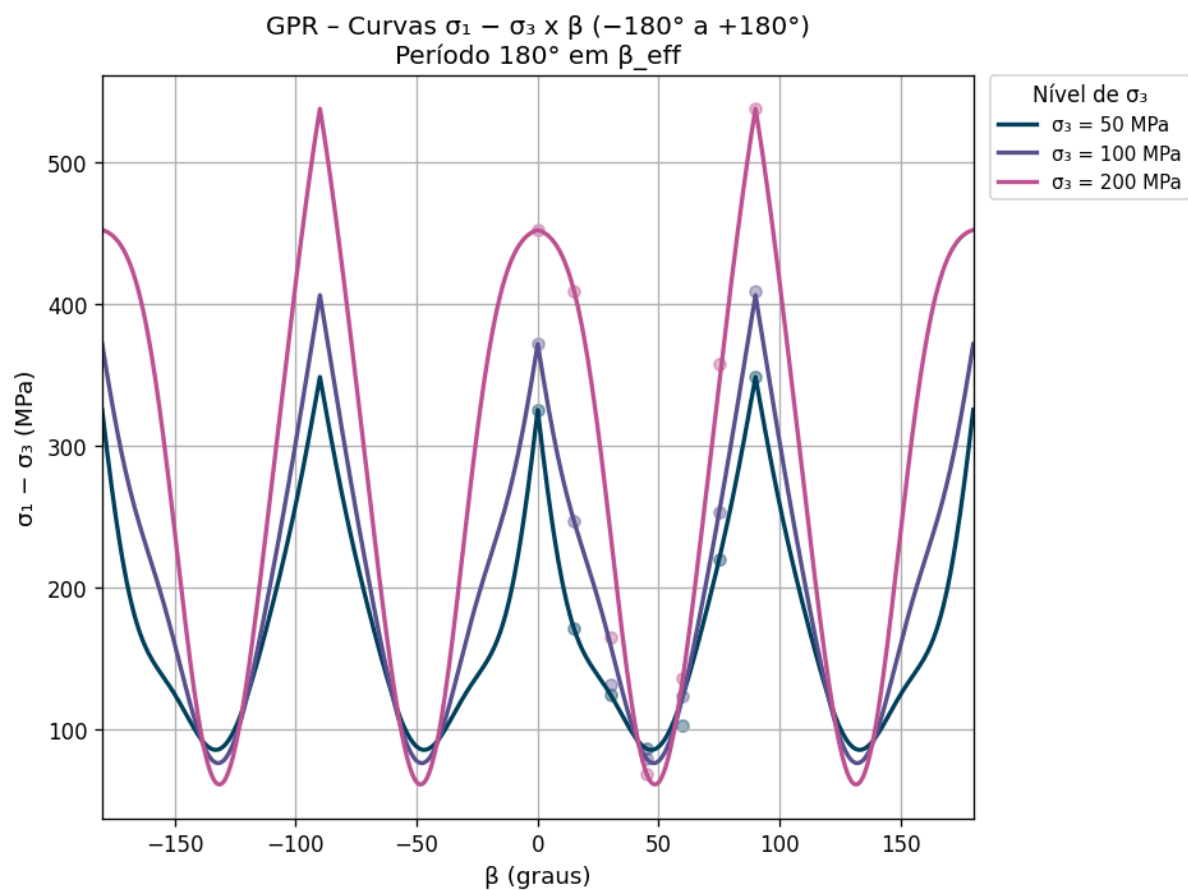
σ_1/σ_3 vs β — GPR (0° – 360° , período 180°)

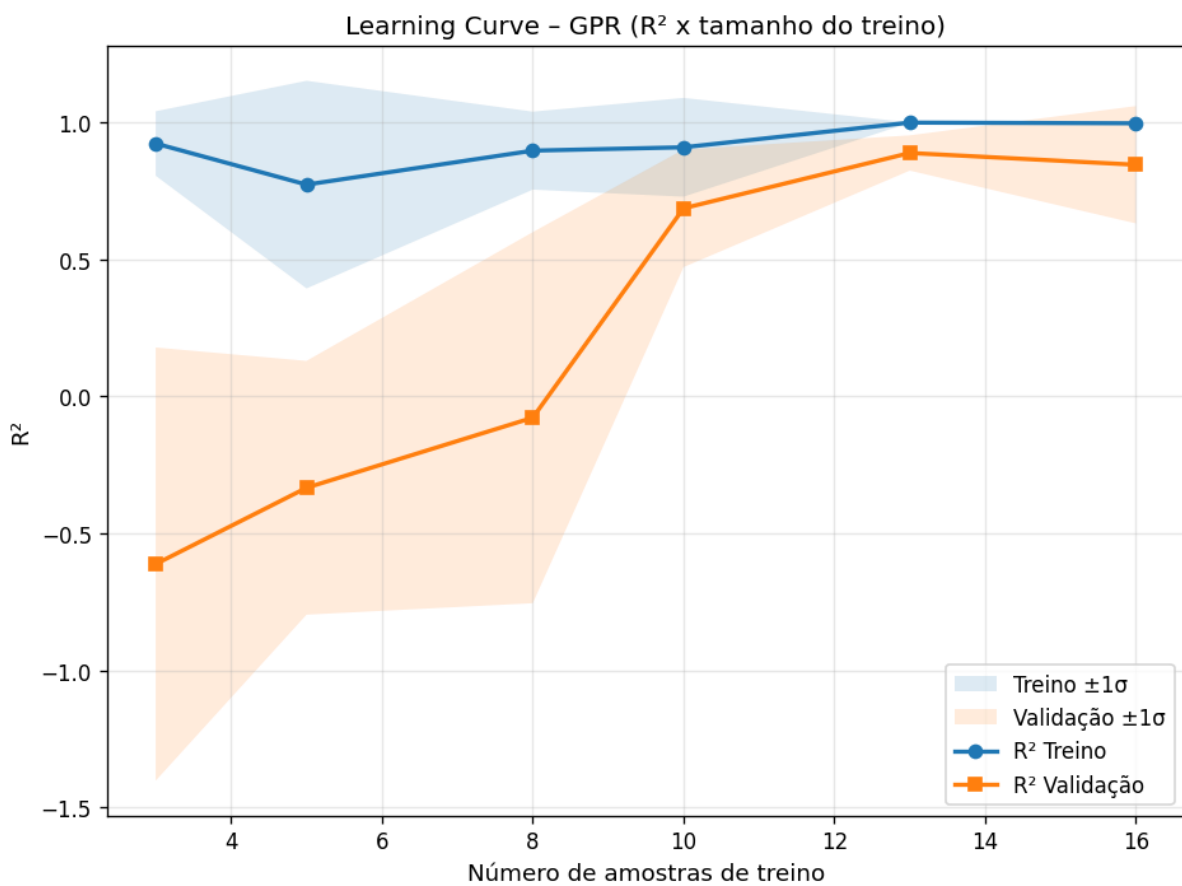
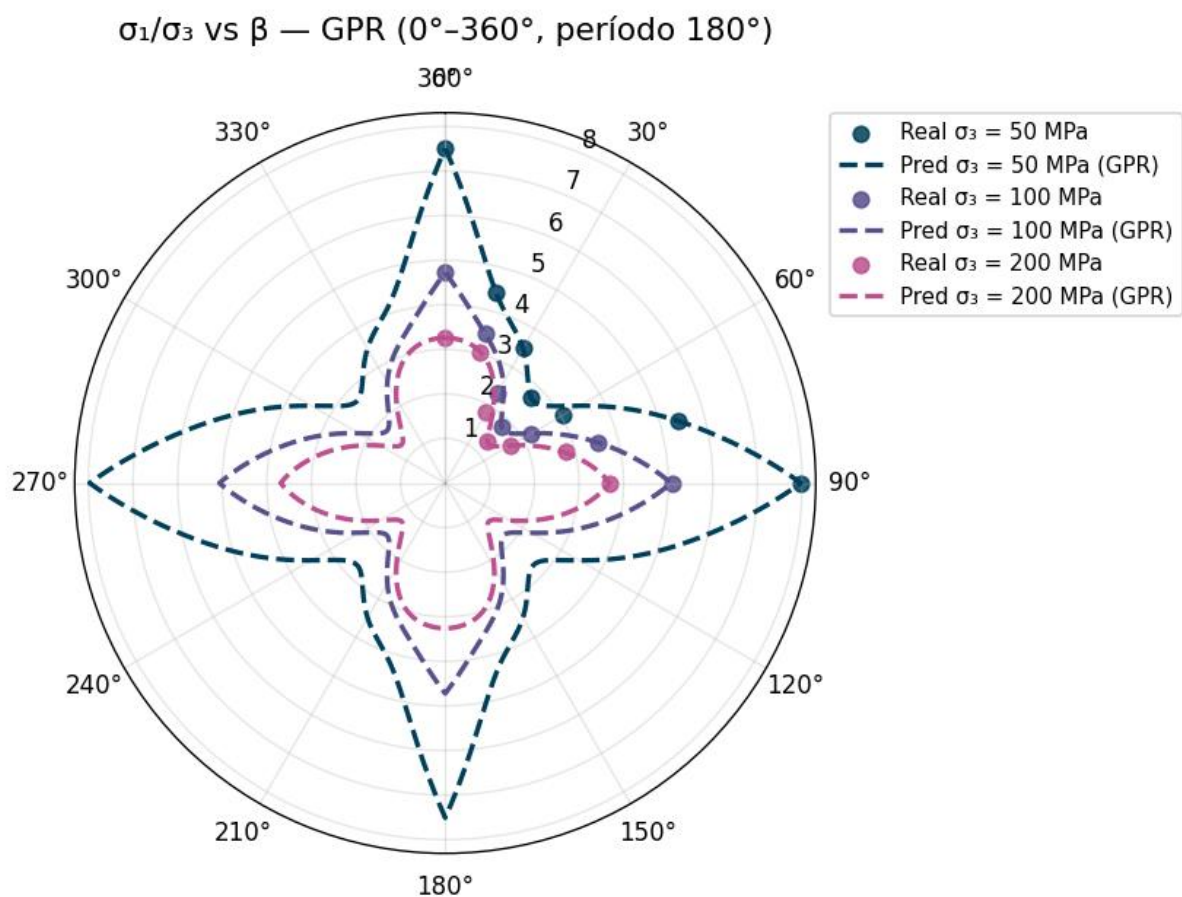




• *Moretown:*

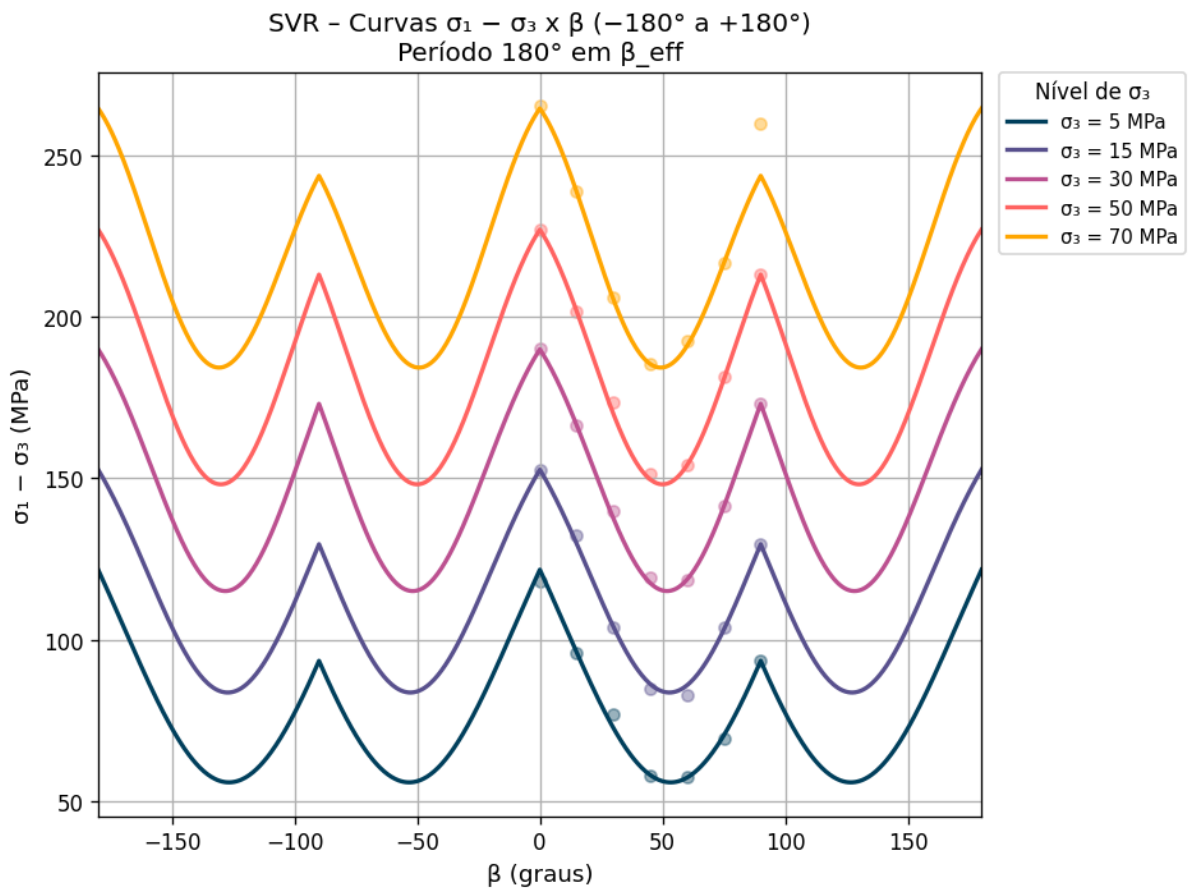
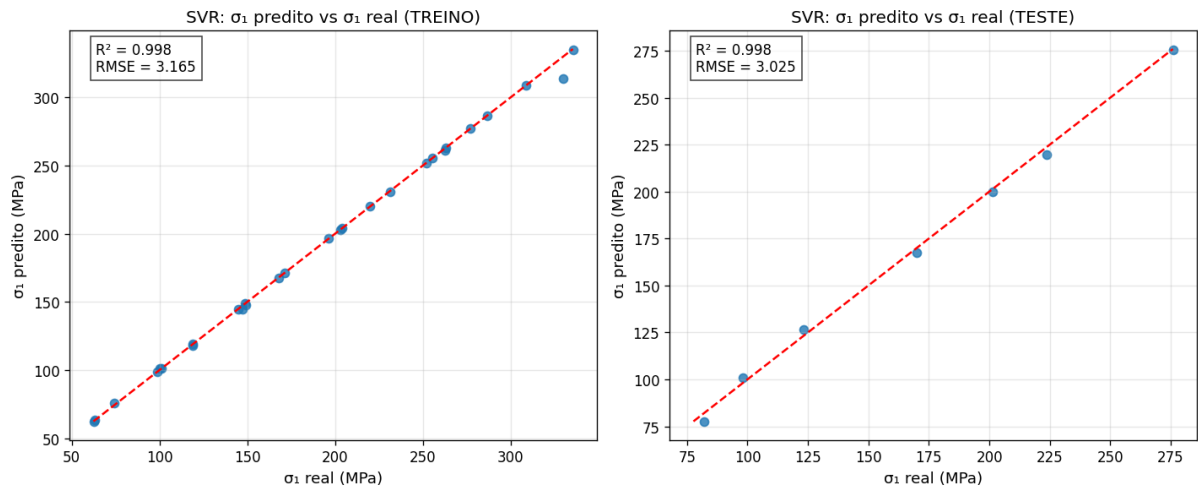




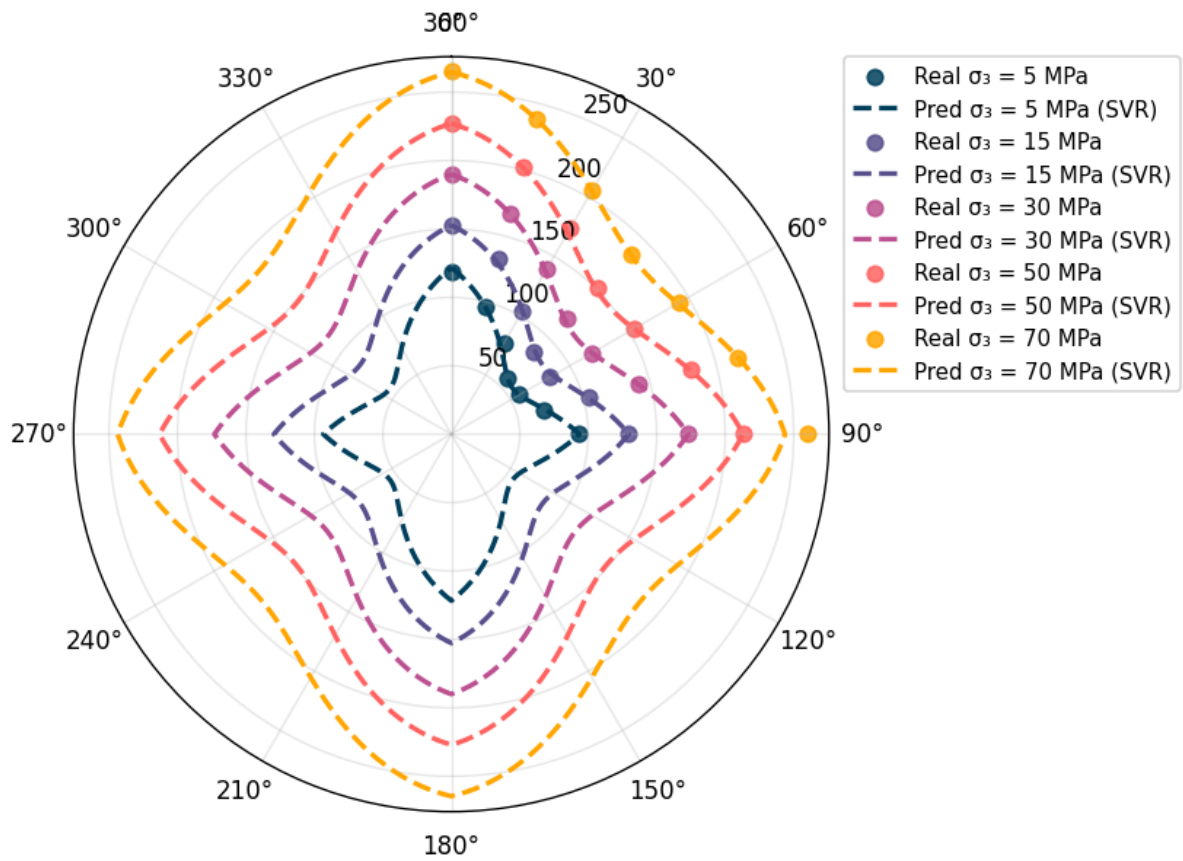


2.4. Resultados do SVR:

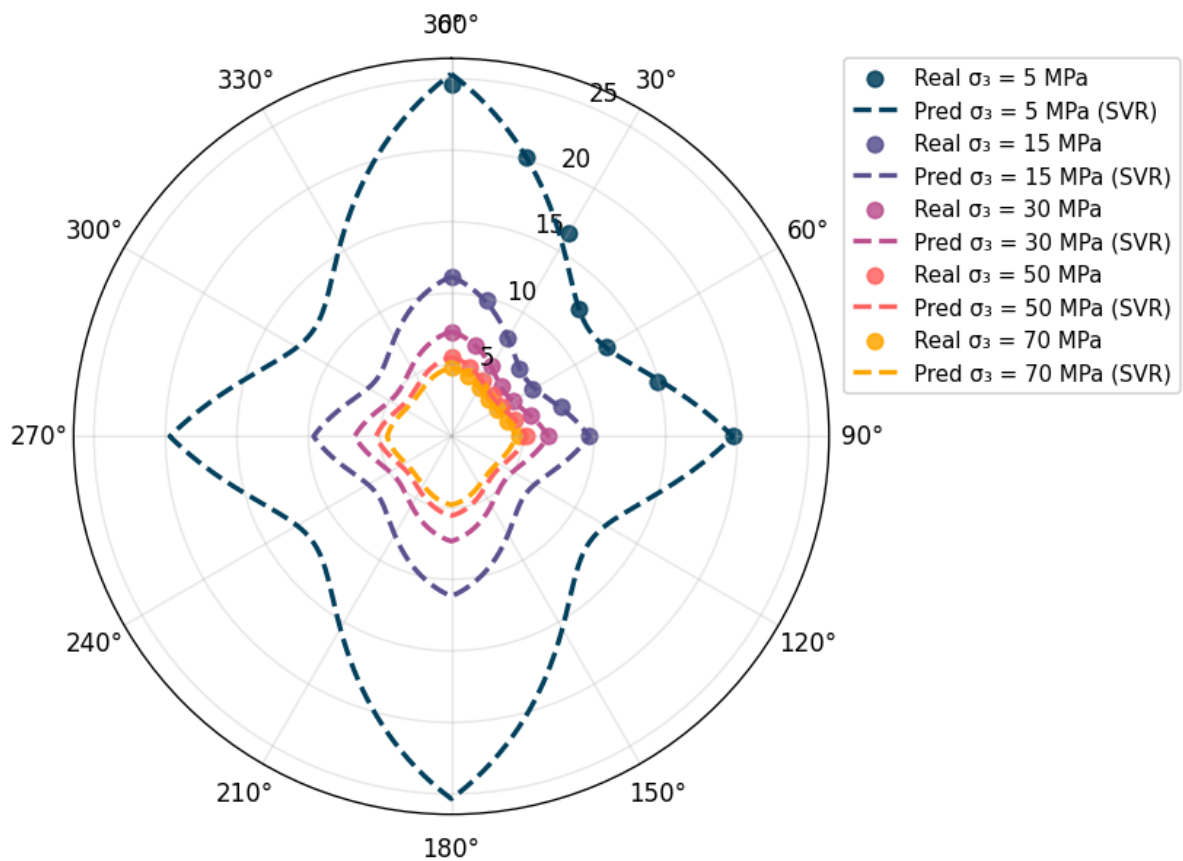
- *Carbonaceous:*

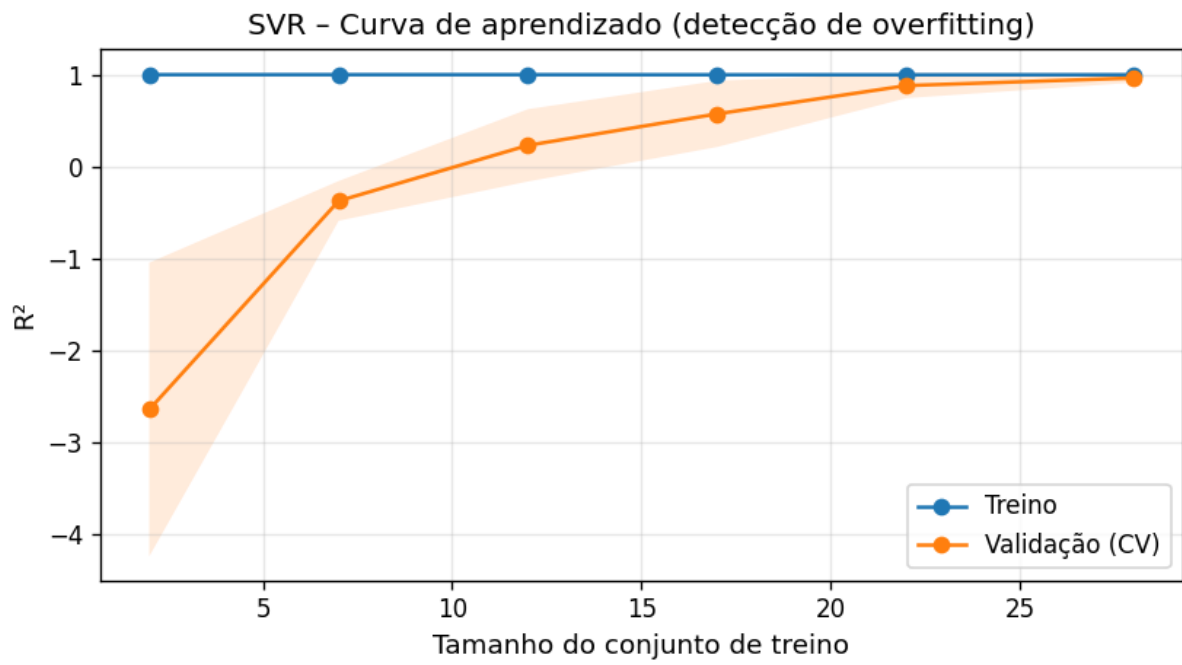


$(\sigma_1 - \sigma_3)$ vs β — SVR (0°-360°, período 180°)

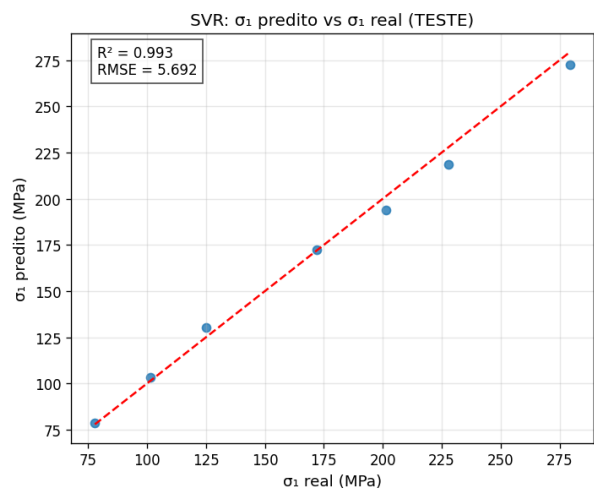
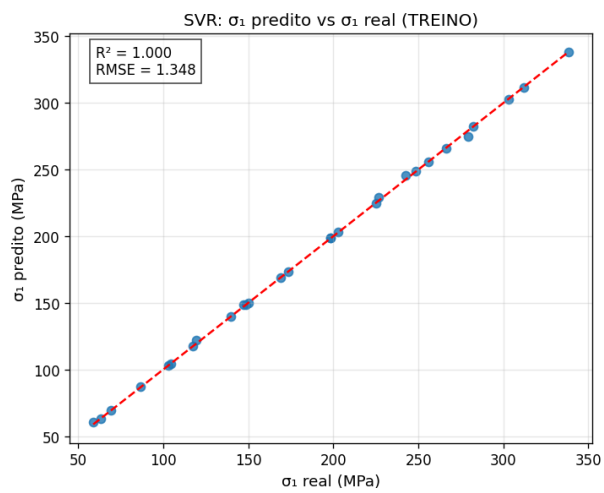


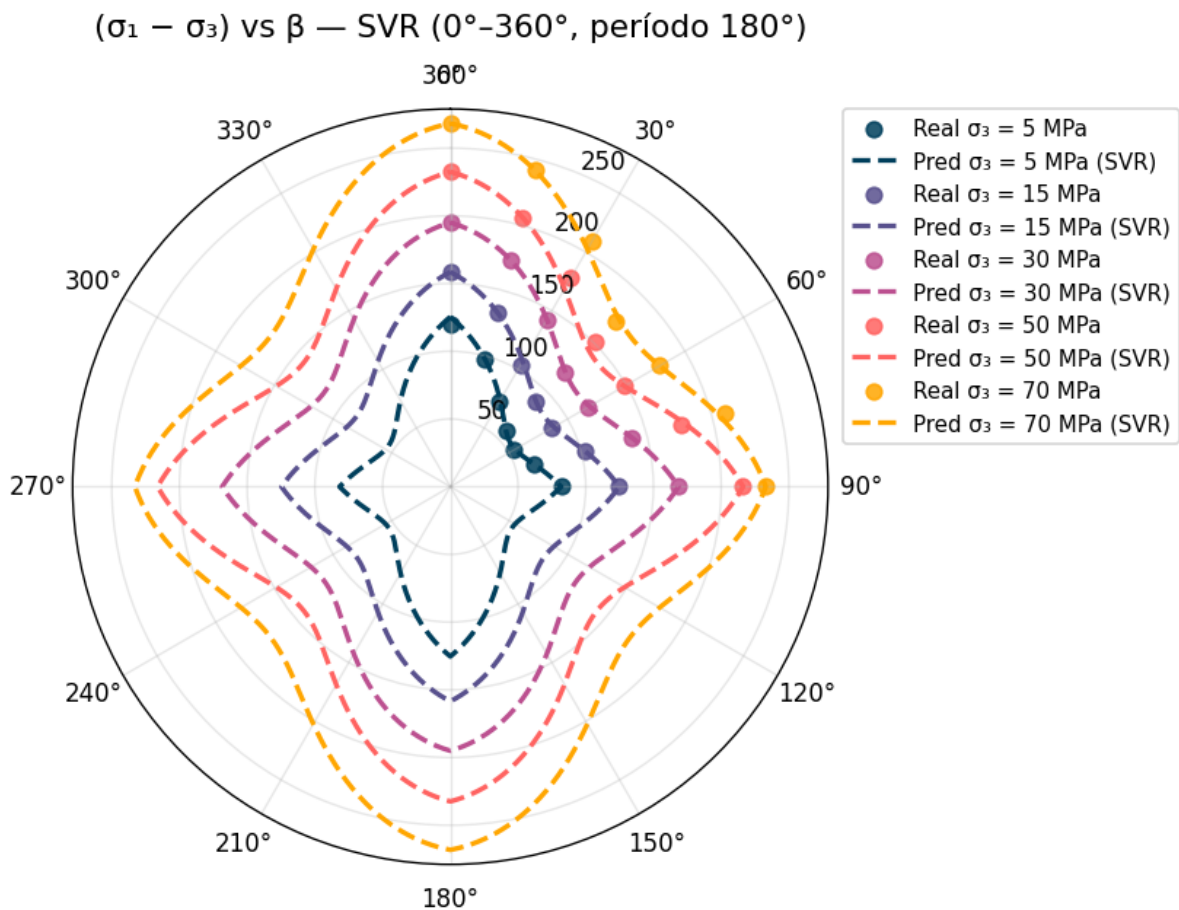
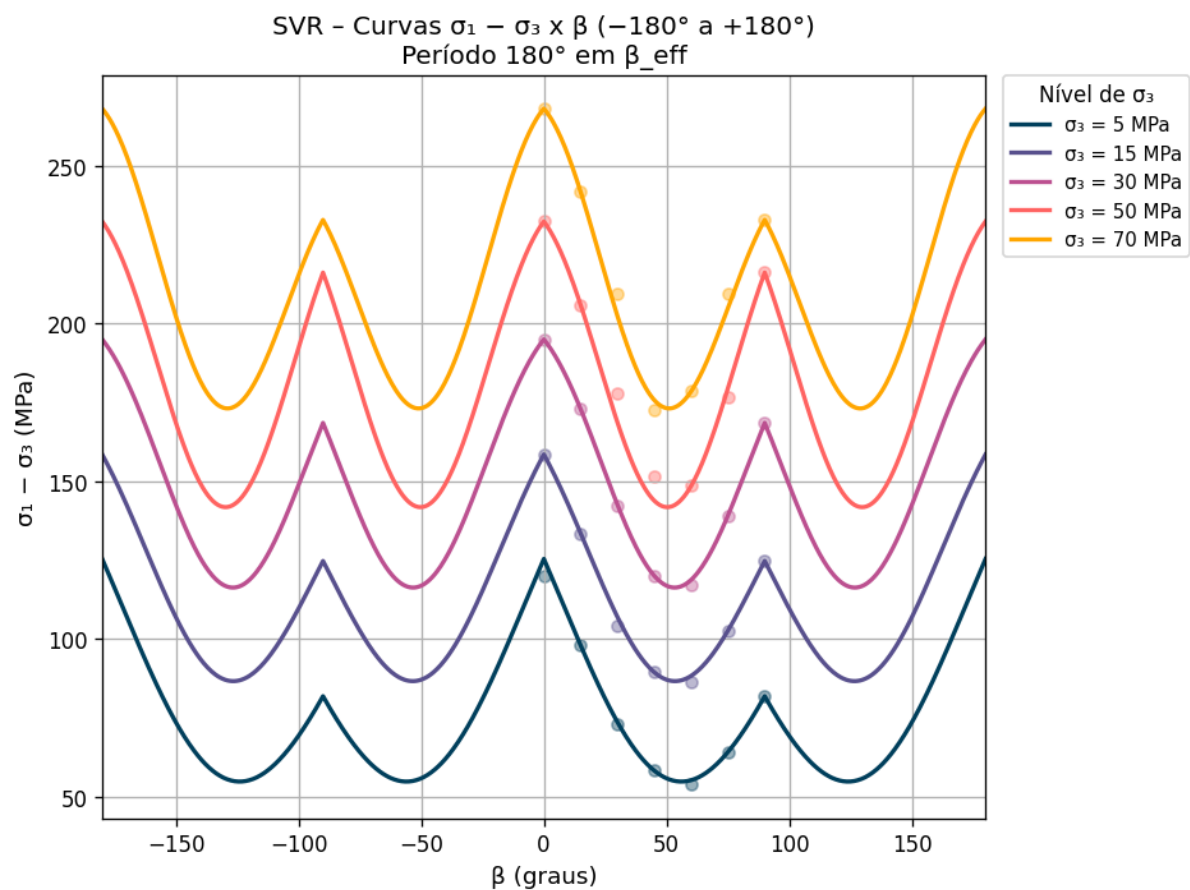
σ_1/σ_3 vs β — SVR (0°-360°, período 180°)

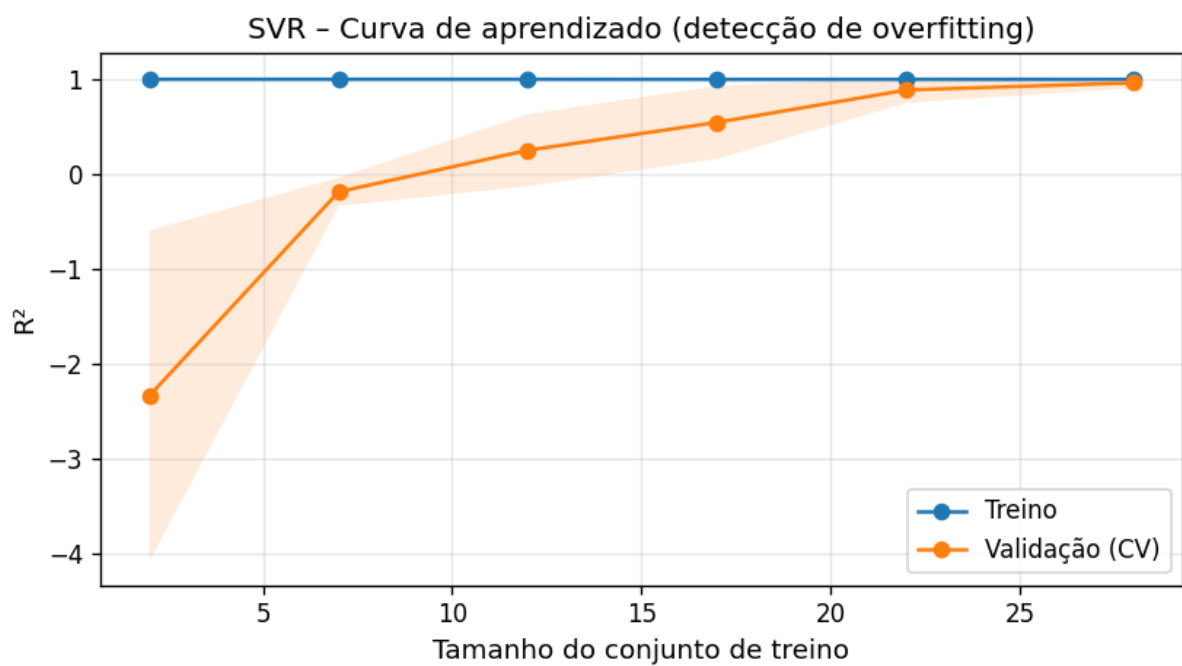
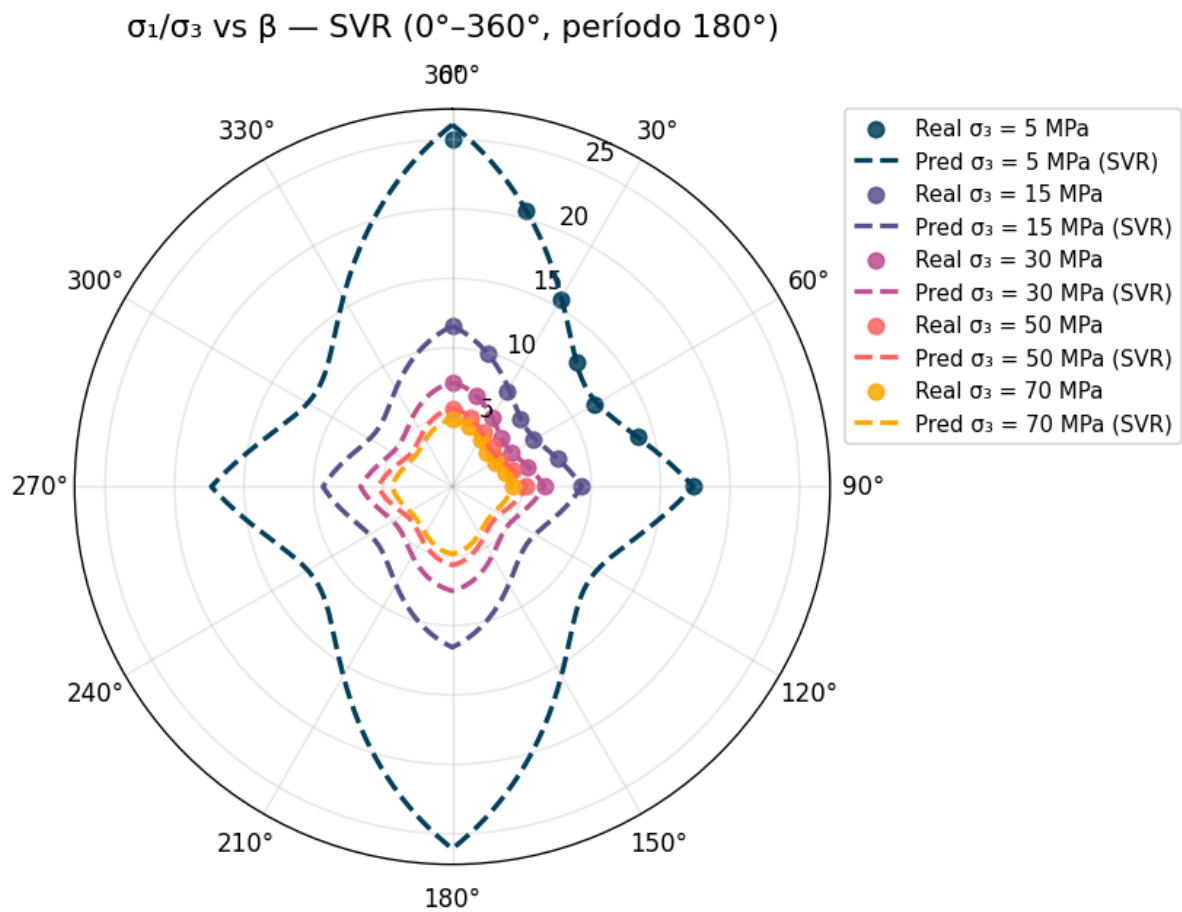




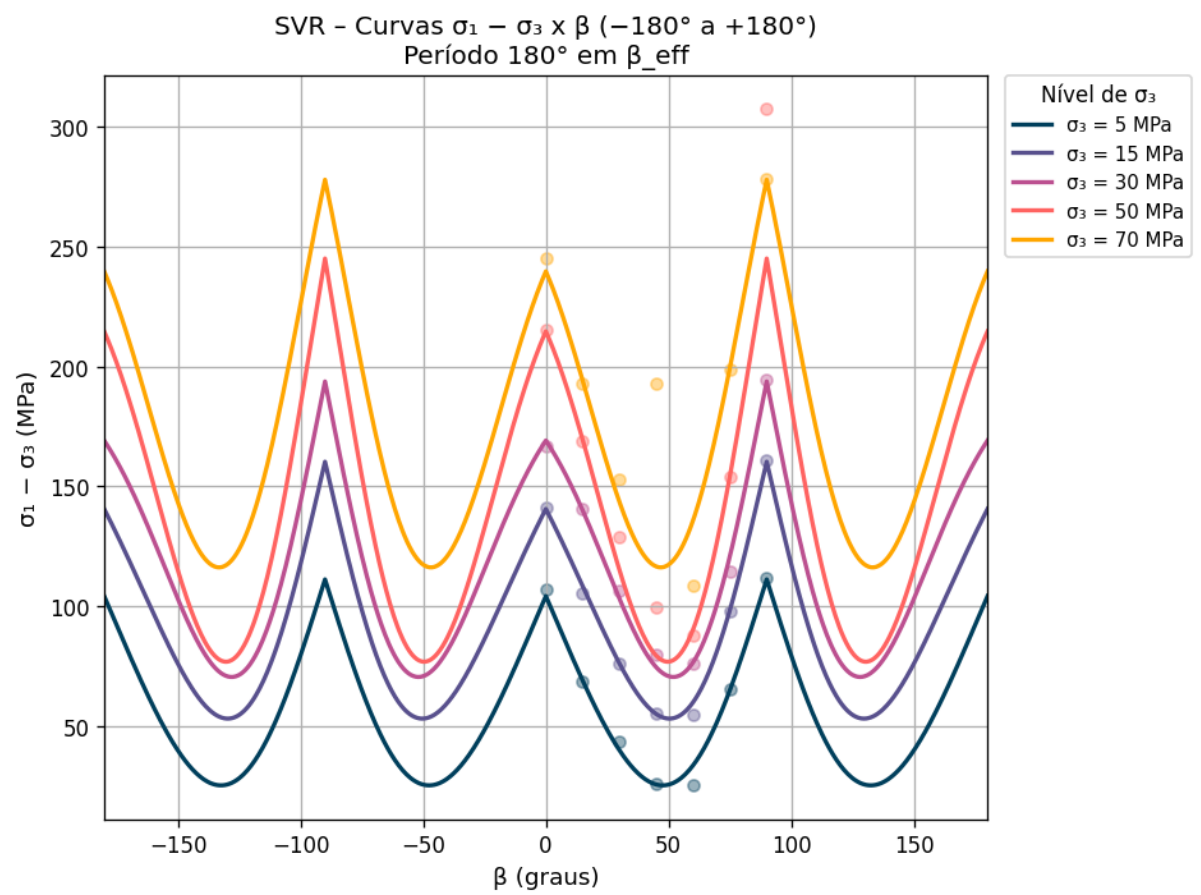
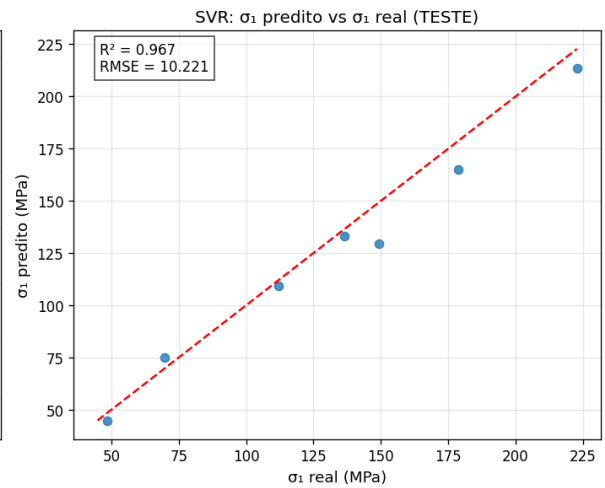
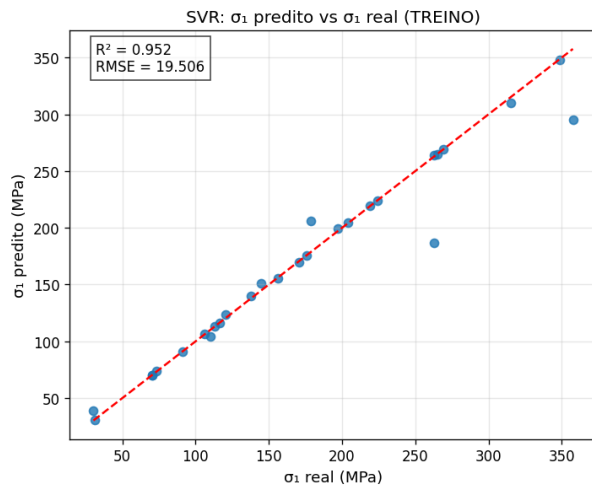
- Quartzite:*



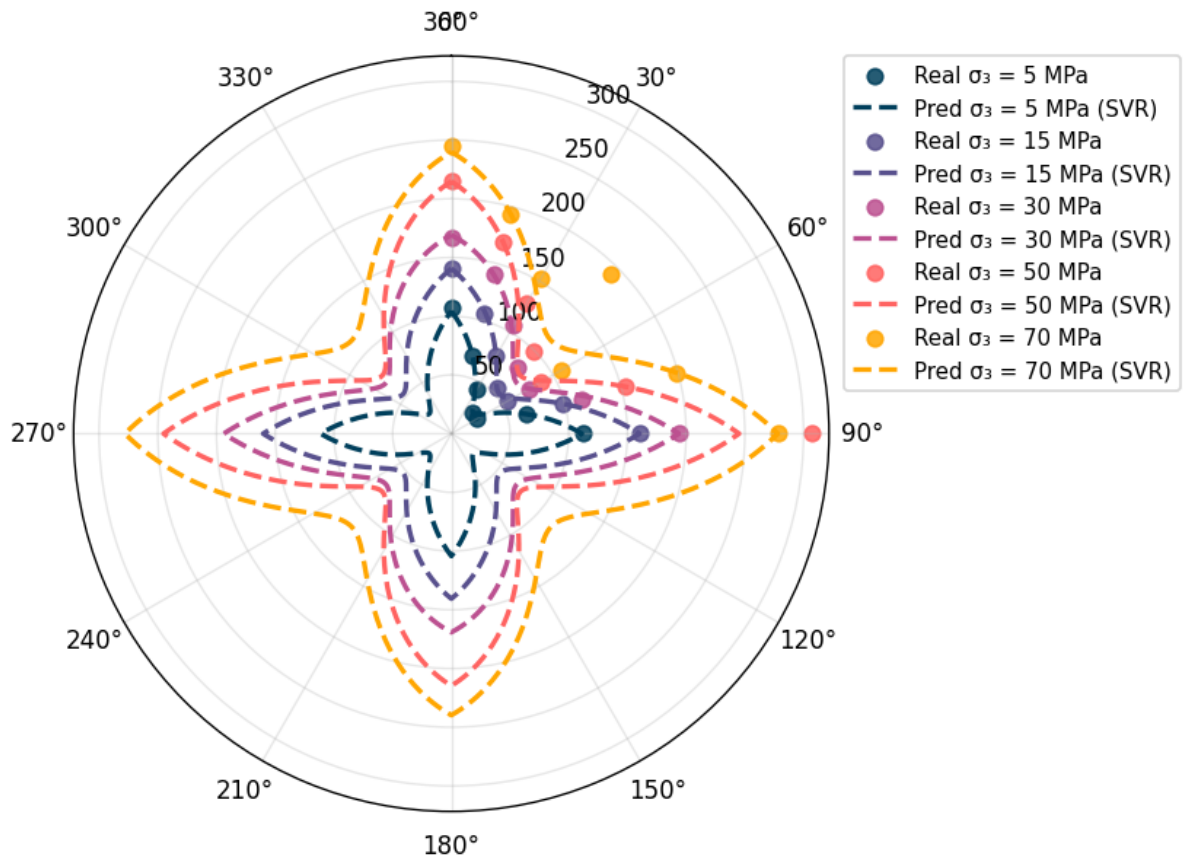




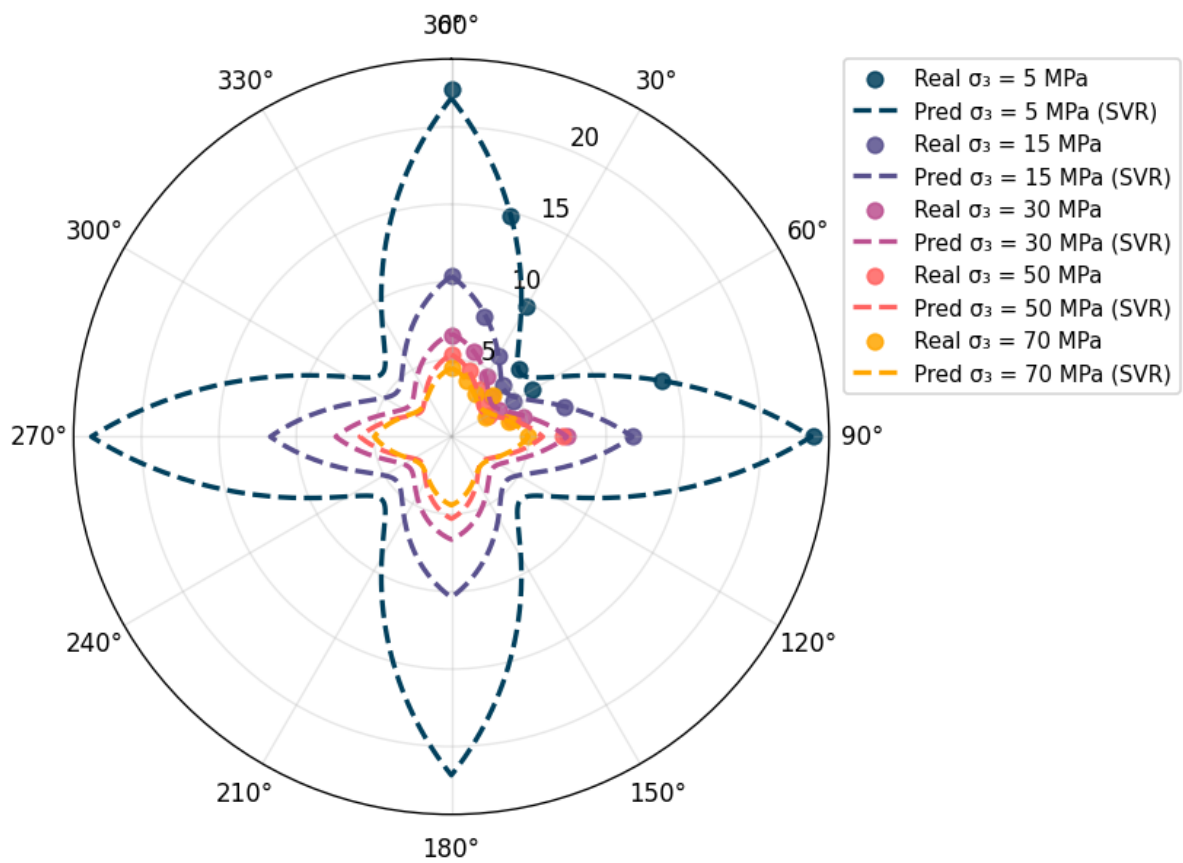
- *Micaceous:*

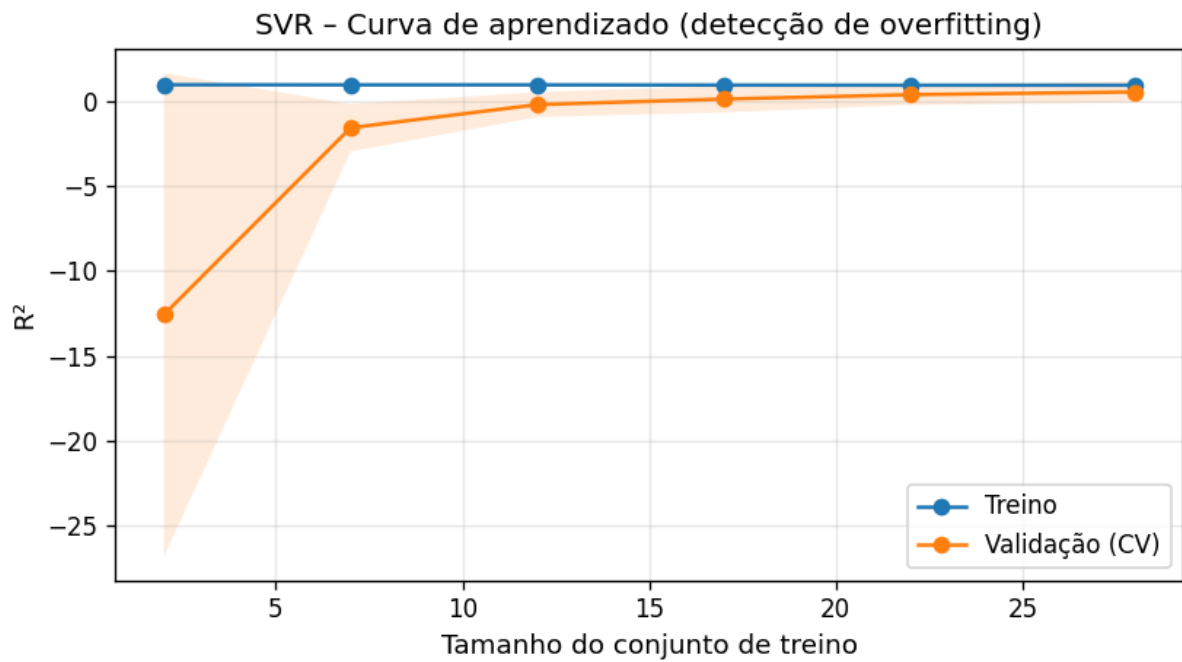


$(\sigma_1 - \sigma_3)$ vs β — SVR (0° – 360° , período 180°)

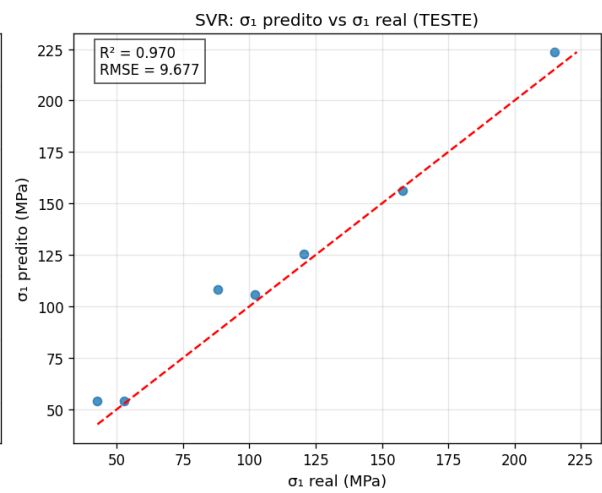
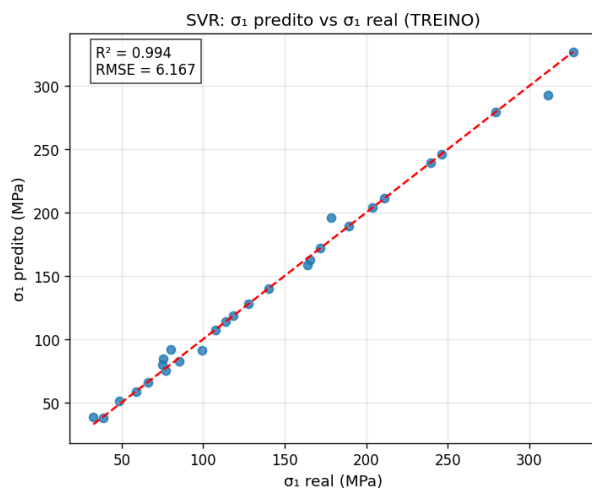


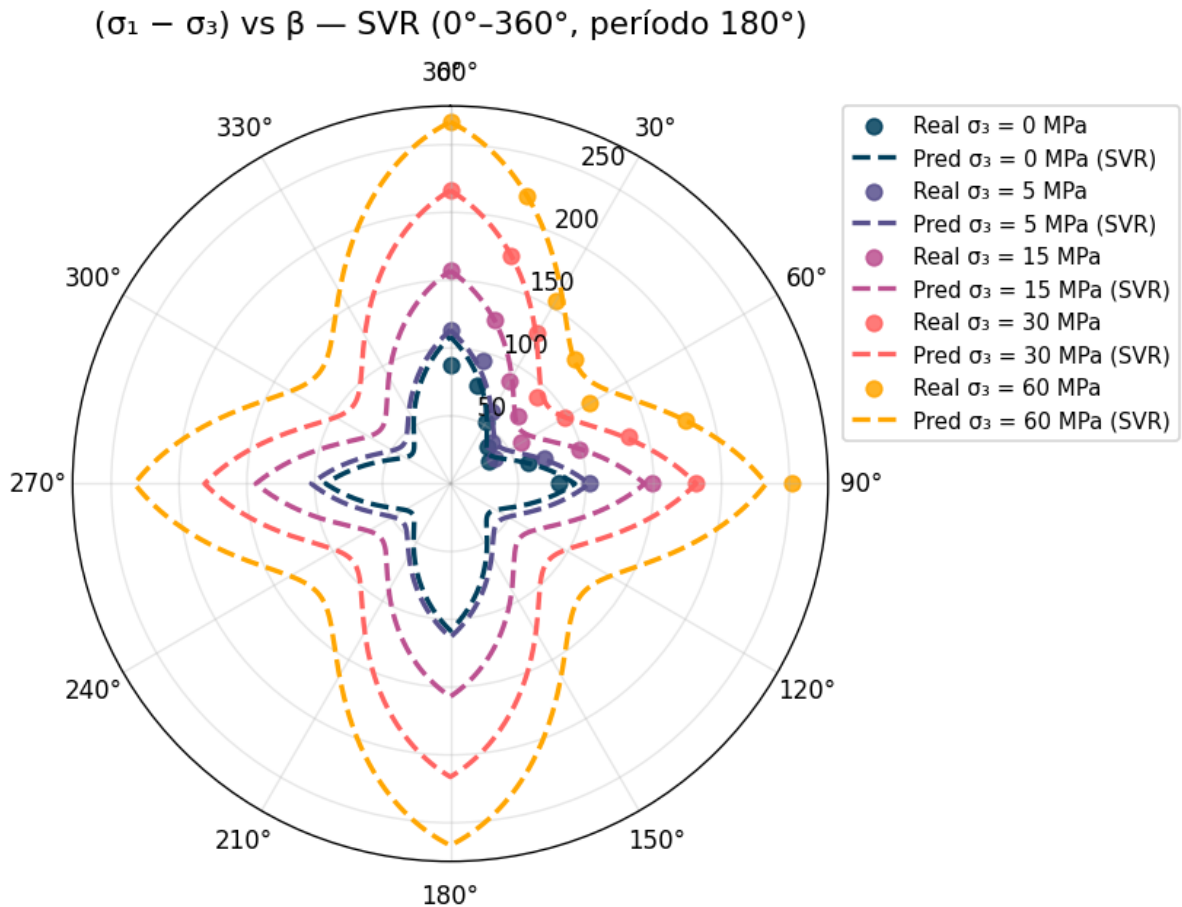
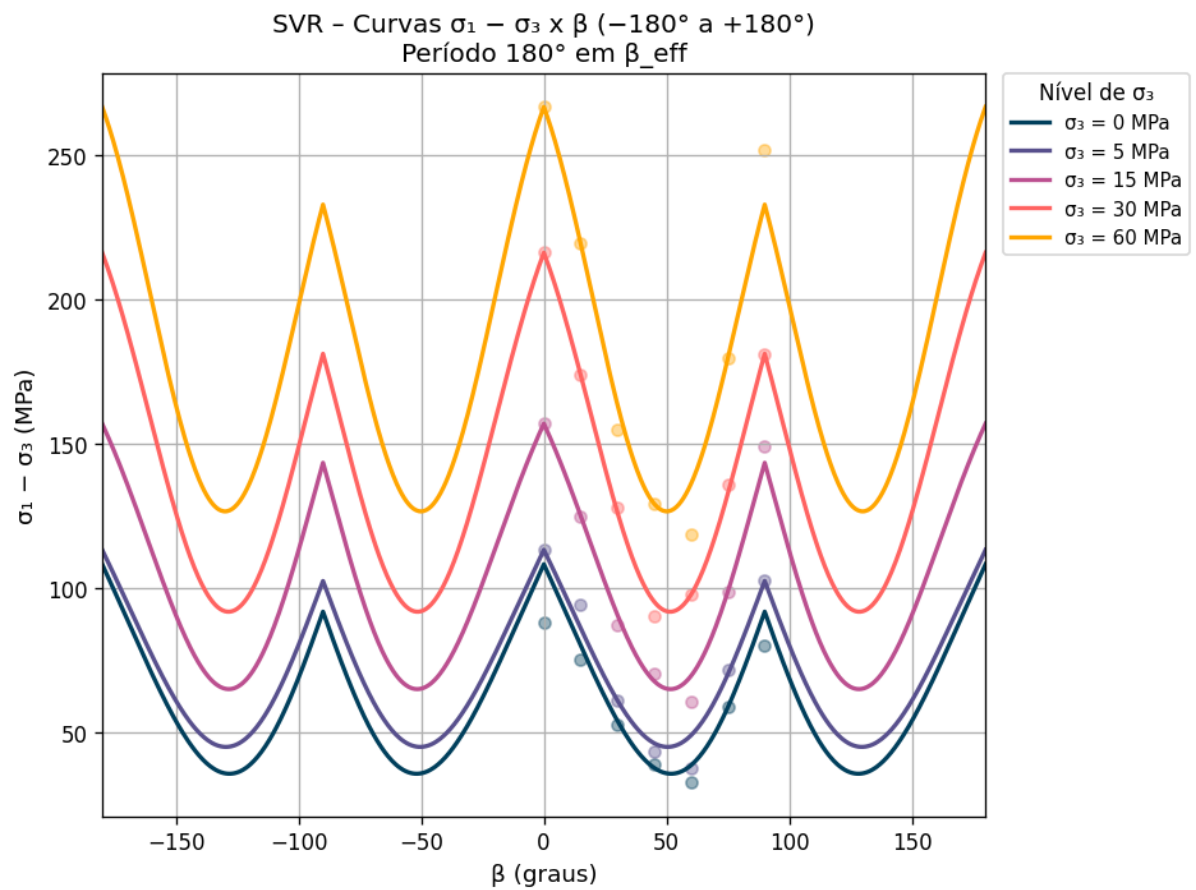
σ_1/σ_3 vs β — SVR (0° – 360° , período 180°)

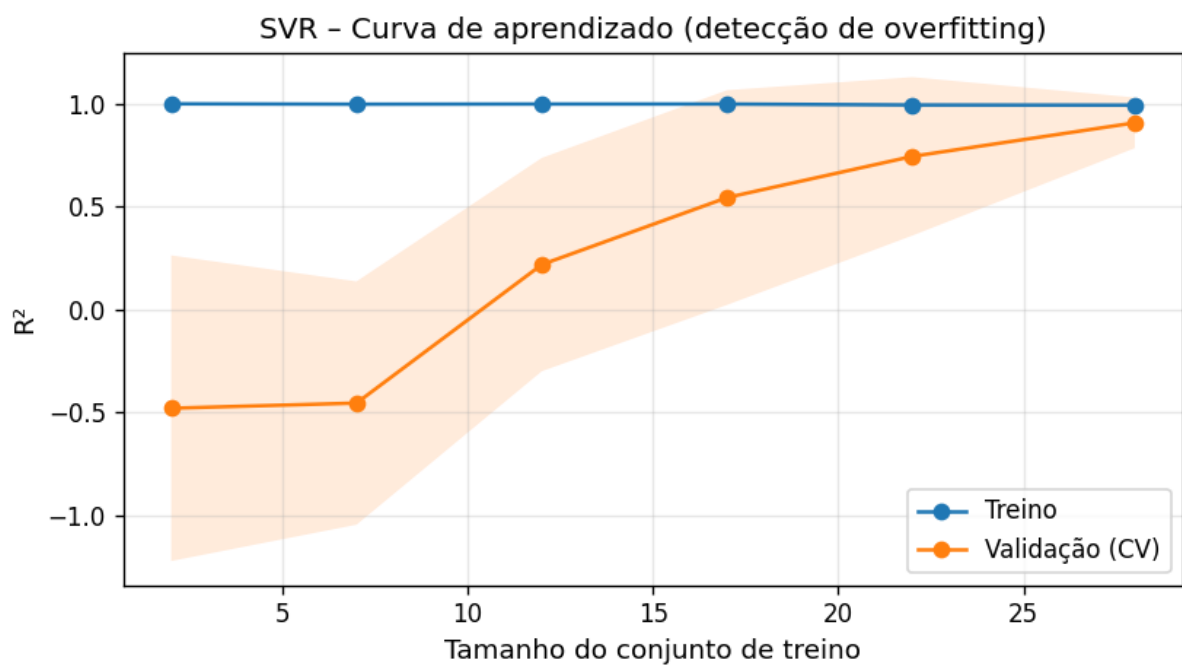
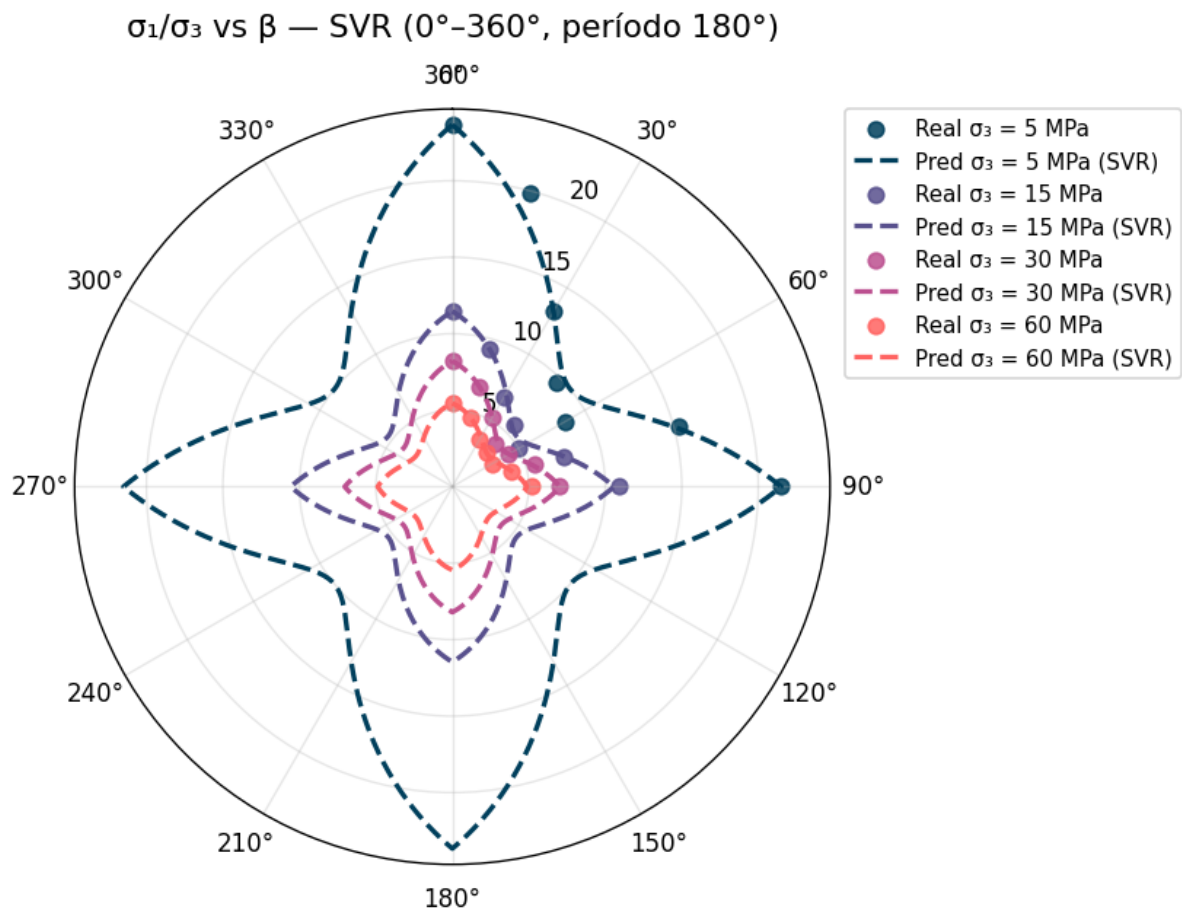




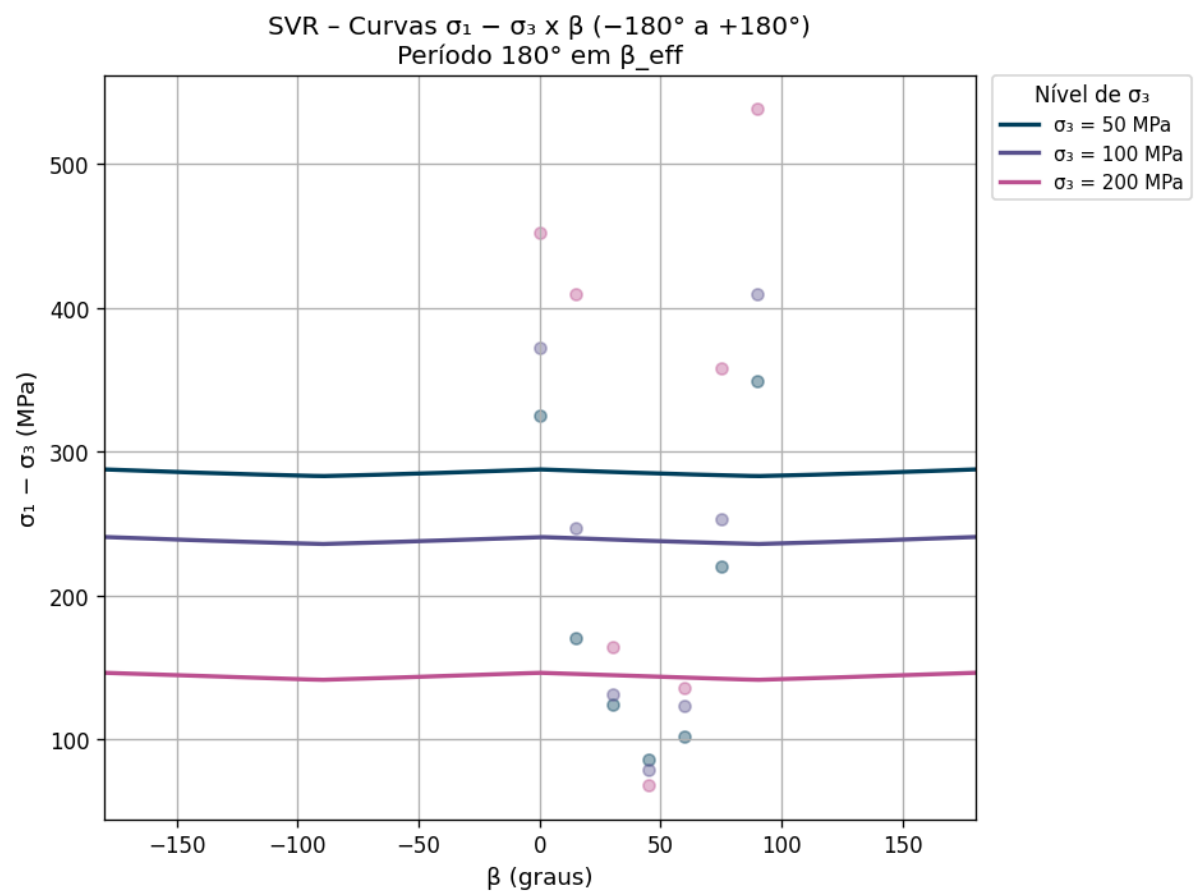
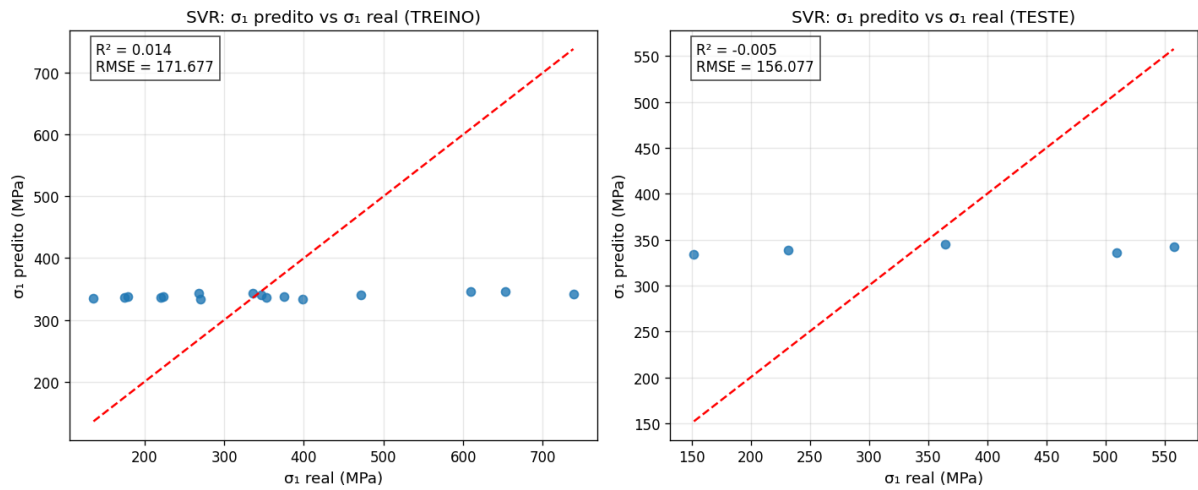
- *Dry:*



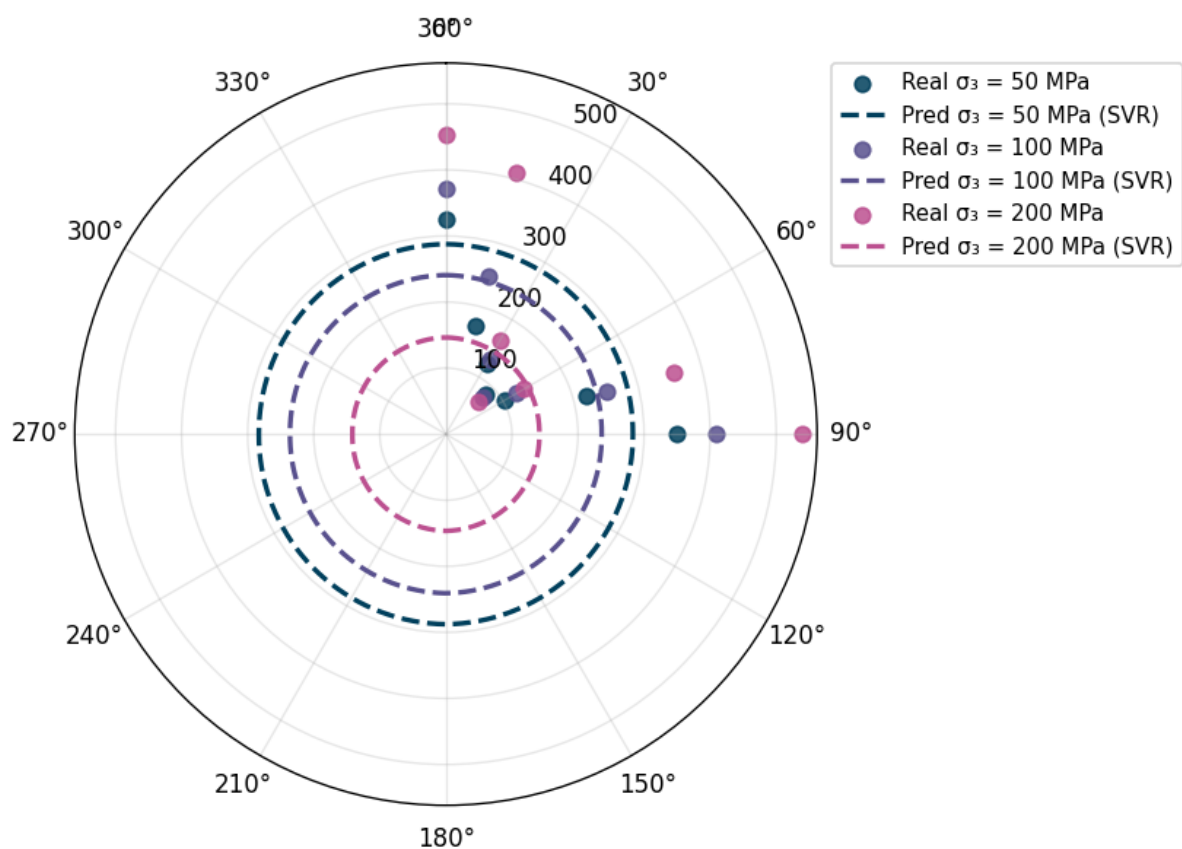




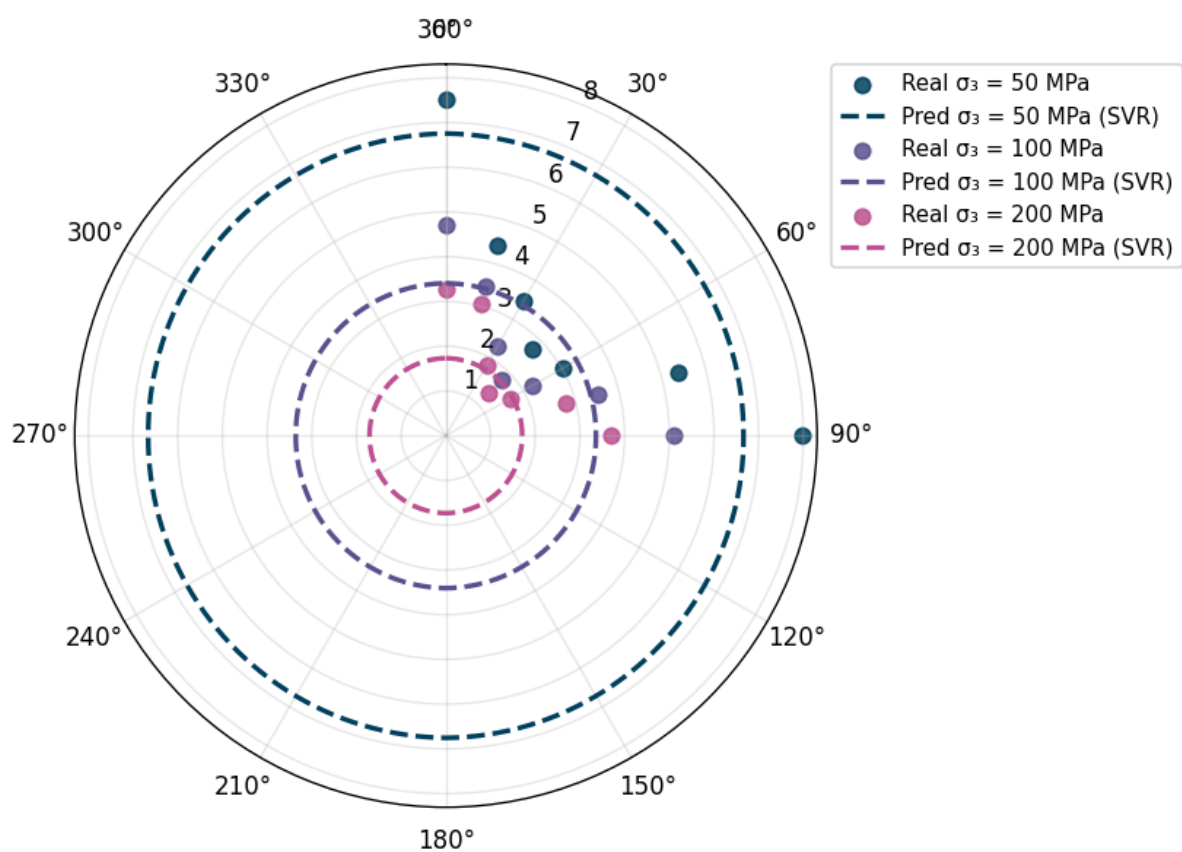
- *Moretown:*

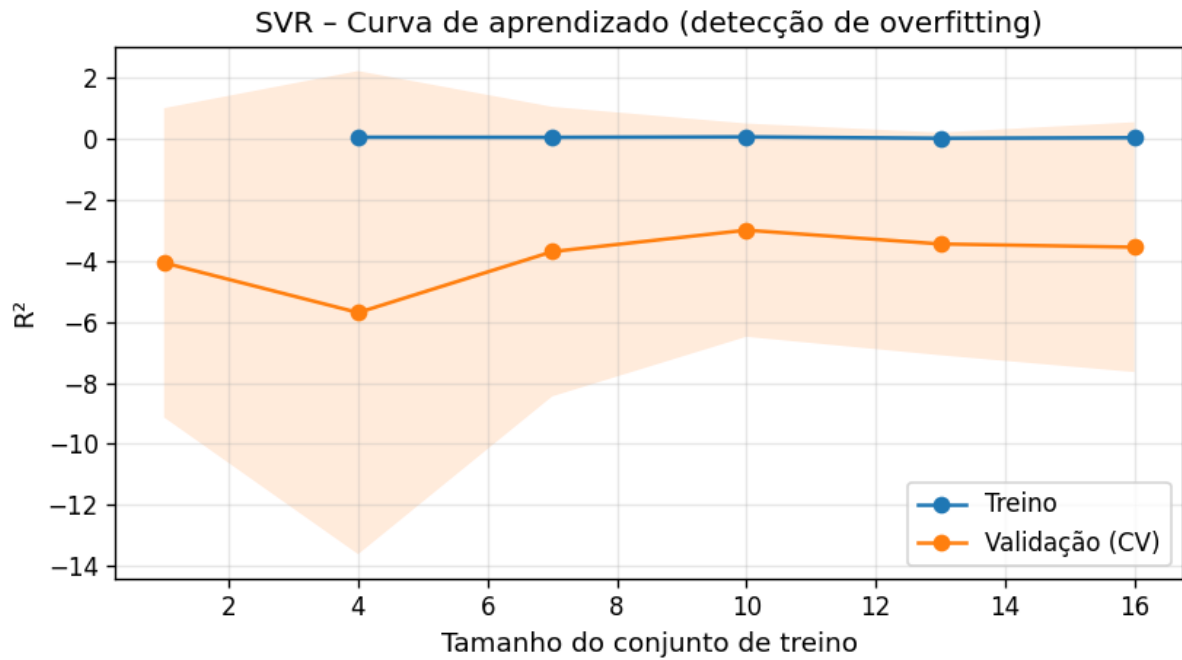


$(\sigma_1 - \sigma_3)$ vs β — SVR (0°-360°, período 180°)



σ_1/σ_3 vs β — SVR (0°-360°, período 180°)





2.5. Estatística descritiva e comparação dos modelos:

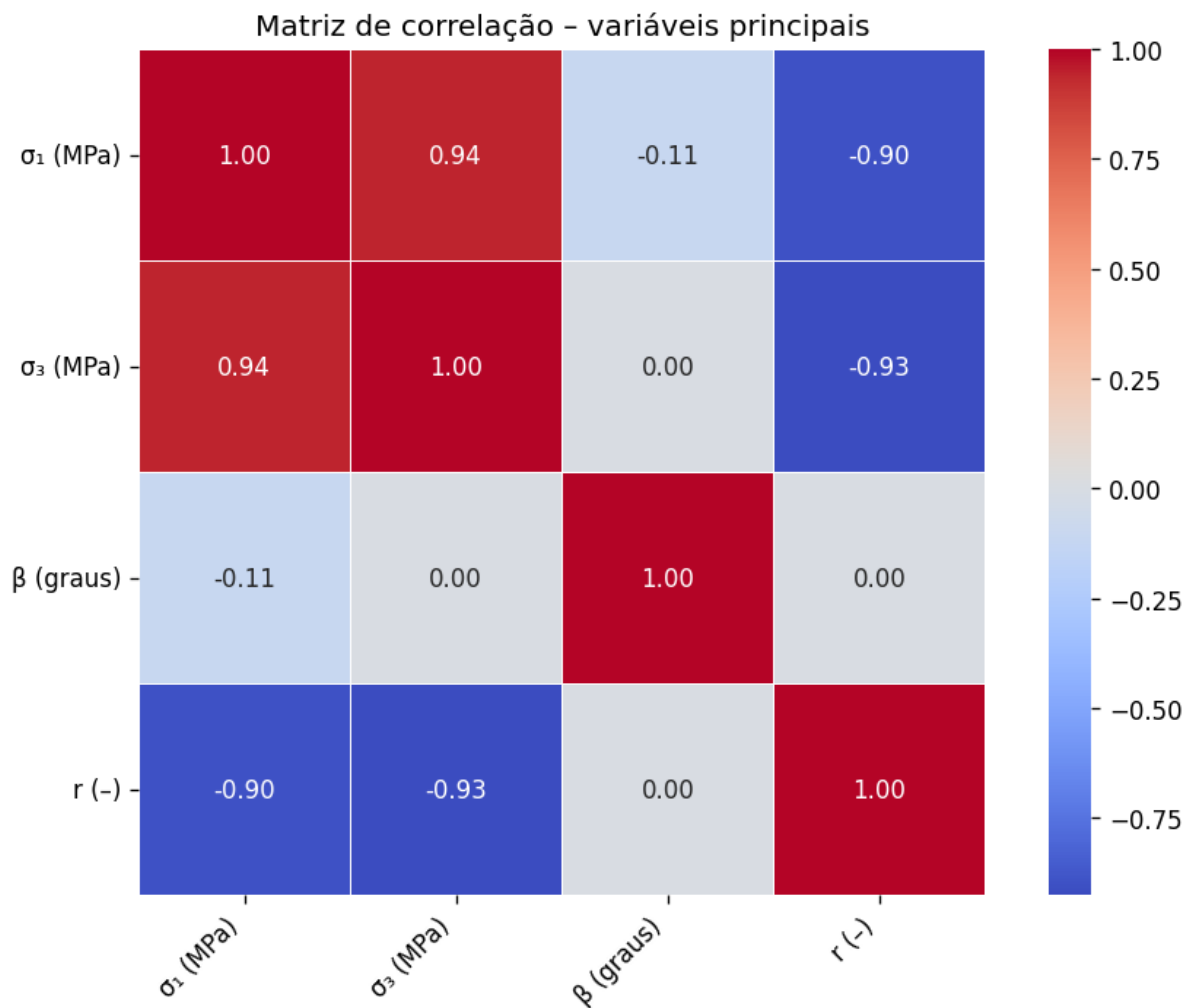
- *Carbonaceous:*

Estatística descritiva global

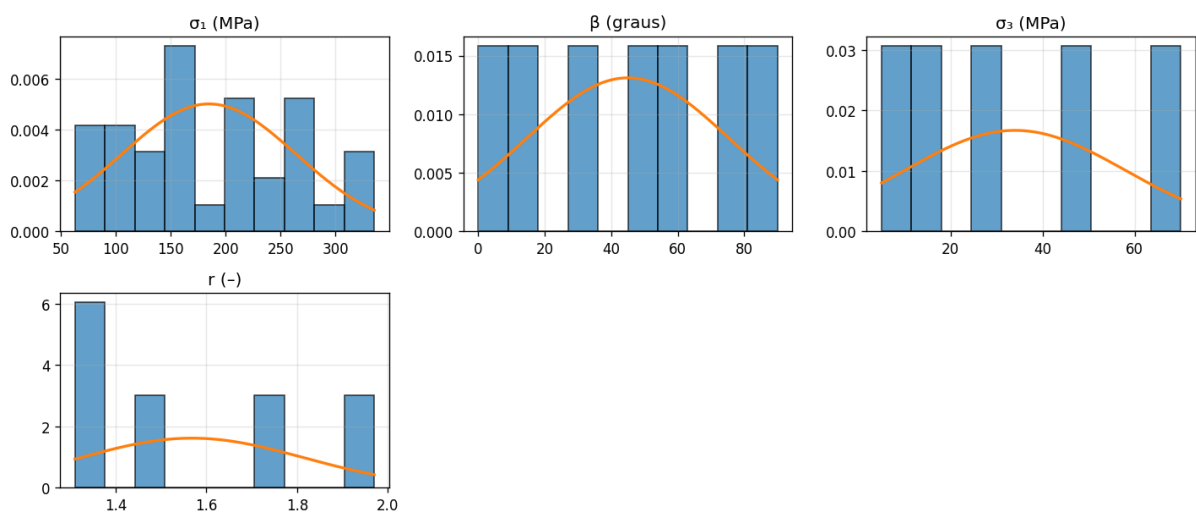
	N	Média	Mediana	Moda	Desvio-padrão	Variância	CV (%)	Mínimo	Q1	Q3	Máximo	Assimetria	Curtose
σ_1 (MPa)	35.00	184.74	171.30	119.00	79.41	6305.54	42.98	62.60	119.00	253.60	335.40	0.20	-1.04
β_{eff} (graus, período 180°)	35.00	45.00	45.00	0.00	30.44	926.47	67.64	0.00	15.00	75.00	90.00	0.00	-1.26
σ_2 (MPa)	35.00	34.00	30.00	5.00	23.88	570.29	70.24	5.00	15.00	50.00	70.00	0.31	-1.33
r (-)	35.00	1.57	1.48	1.31	0.25	0.06	15.72	1.31	1.37	1.71	1.97	0.61	-1.12
σ_c (MPa)	35.00	56.50	56.50	56.50	0.00	0.00	0.00	56.50	56.50	56.50	56.50	0.00	0.00
Coesão (MPa)	35.00	22.00	22.00	22.00	0.00	0.00	0.00	22.00	22.00	22.00	22.00	0.00	0.00
Ângulo de atrito (°)	35.00	31.10	31.10	31.10	0.00	0.00	0.00	31.10	31.10	31.10	31.10	0.00	0.00
$\sigma_1 - \sigma_2$ (MPa)	35.00	150.74	151.50	104.00	57.43	3298.46	38.10	57.60	104.00	191.30	265.40	0.20	-0.83
σ_1/σ_2 (-)	35.00	8.14	5.71	7.93	5.30	28.09	65.08	3.65	4.67	9.73	24.62	1.68	2.22

Estatísticas por nível de σ_2 - Variáveis de resistência

Variável	N	σ_1 (MPa)						$\sigma_1 - \sigma_2$ (MPa)						σ_1/σ_2 (-)					
Estatística		Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo
σ_2 (MPa)																			
5	7	86.41	82.10	22.29	496.87	62.60	123.10	81.41	77.10	22.29	496.87	57.60	118.10	17.28	16.42	4.46	19.87	12.52	24.62
15	7	127.93	119.00	26.03	677.71	98.00	167.70	112.93	104.00	26.03	677.71	83.00	152.70	8.53	7.93	1.74	3.01	6.53	11.18
30	7	179.79	171.30	27.42	751.72	148.70	220.00	149.79	141.30	27.42	751.72	118.70	190.00	5.99	5.71	0.91	0.84	4.96	7.33
50	7	236.04	231.30	28.98	840.09	201.50	276.90	186.04	181.30	28.98	840.09	151.50	226.90	4.72	4.63	0.58	0.34	4.03	5.54
70	7	293.53	286.70	31.84	1013.55	255.40	335.40	223.53	216.70	31.84	1013.55	185.40	265.40	4.19	4.10	0.45	0.21	3.65	4.79



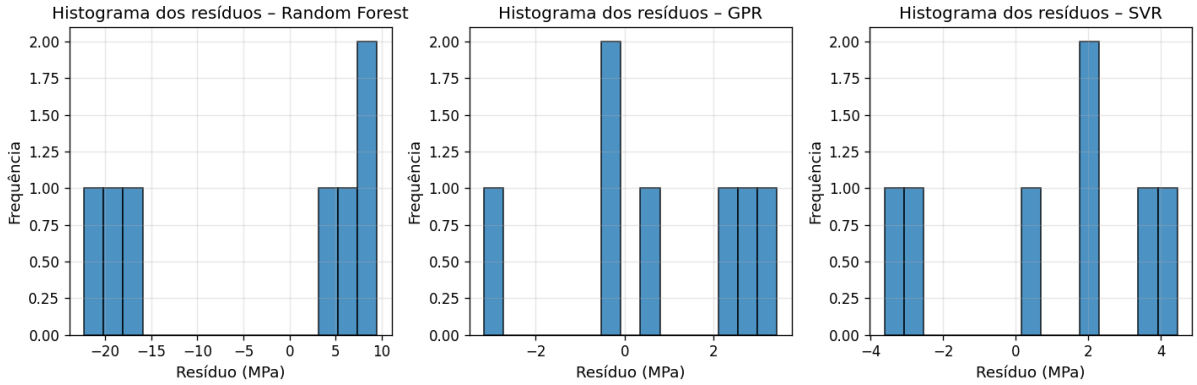
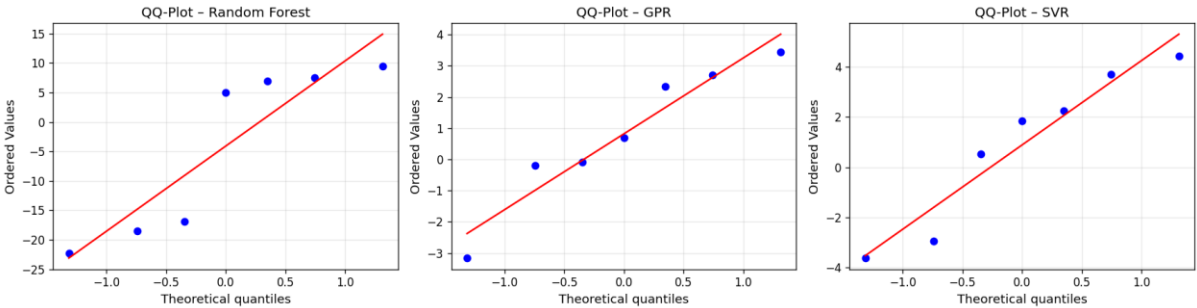
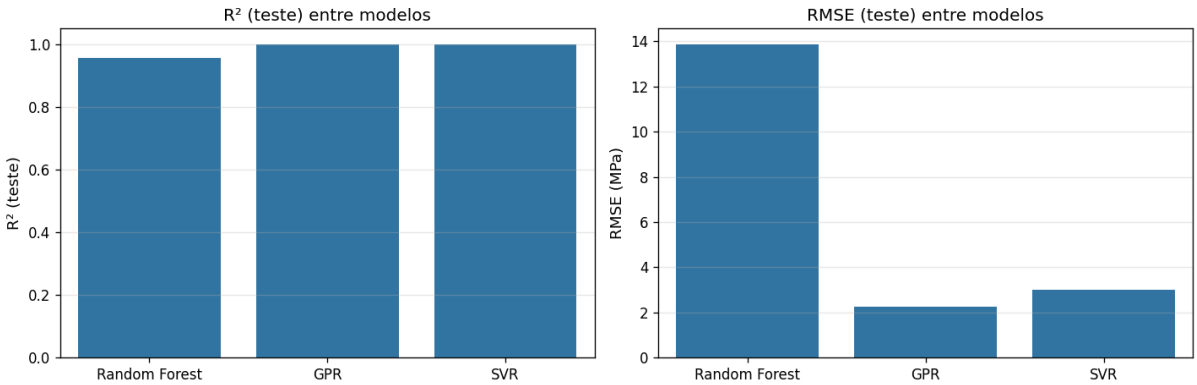
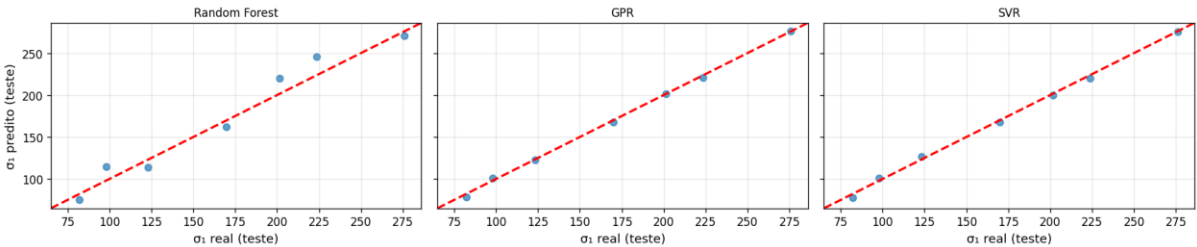
Histogramas + curva Normal - variáveis base

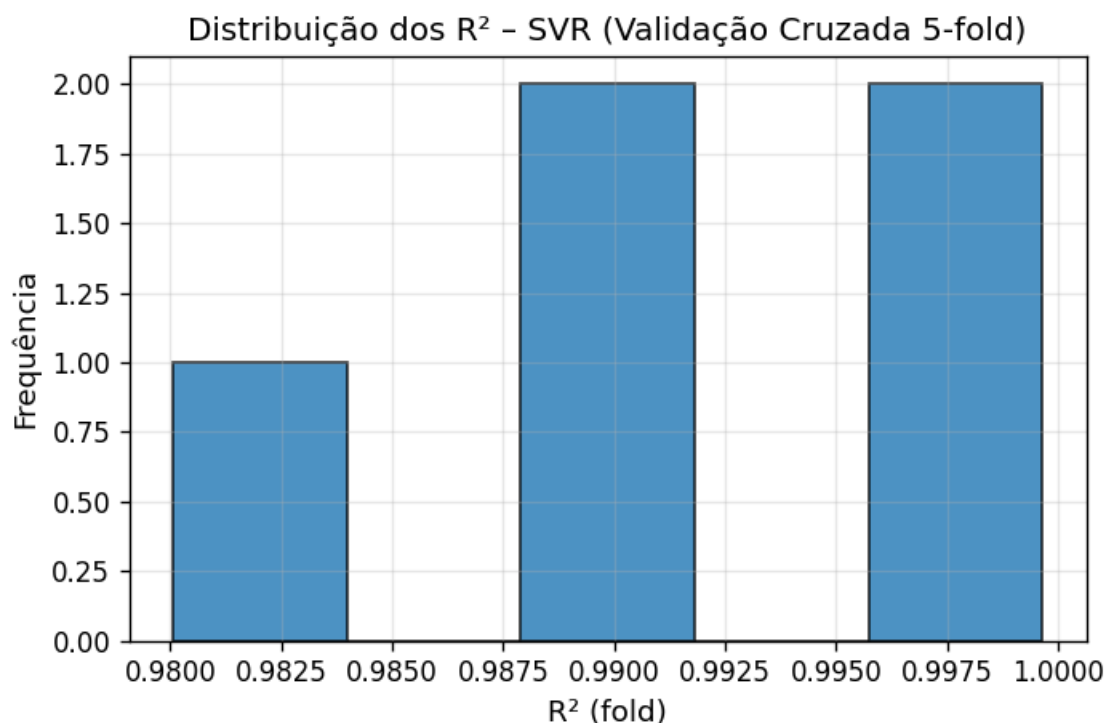
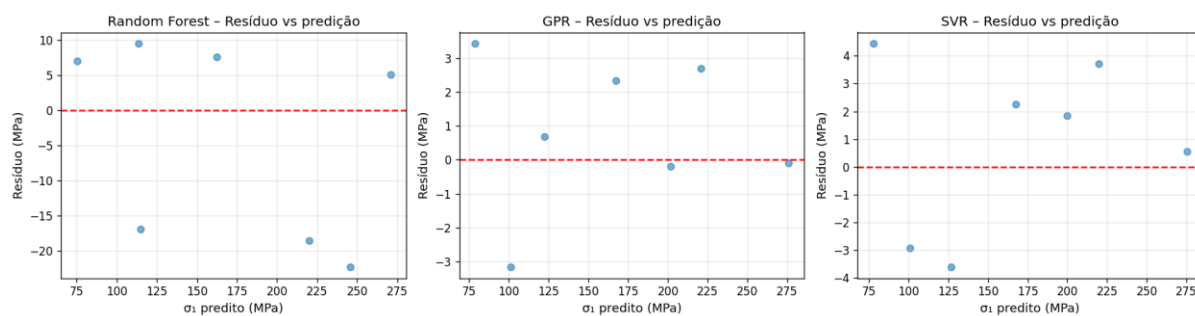


Métricas globais dos modelos (incluindo validação cruzada)

	Modelo	R2_treino	R2_teste	RMSE_treino	RMSE_teste	MAE_treino	MAE_teste	R2_CV_medio	R2_CV_std
0	Random Forest	0.99	0.96	7.58	13.88	6.15	12.40	0.87	0.14
1	GPR	1.00	1.00	1.16	2.24	0.94	1.80	0.99	0.01
2	SVR	1.00	1.00	3.17	3.03	0.97	2.76	0.96	0.05

Comparação visual - σ_1 real vs predito (teste)





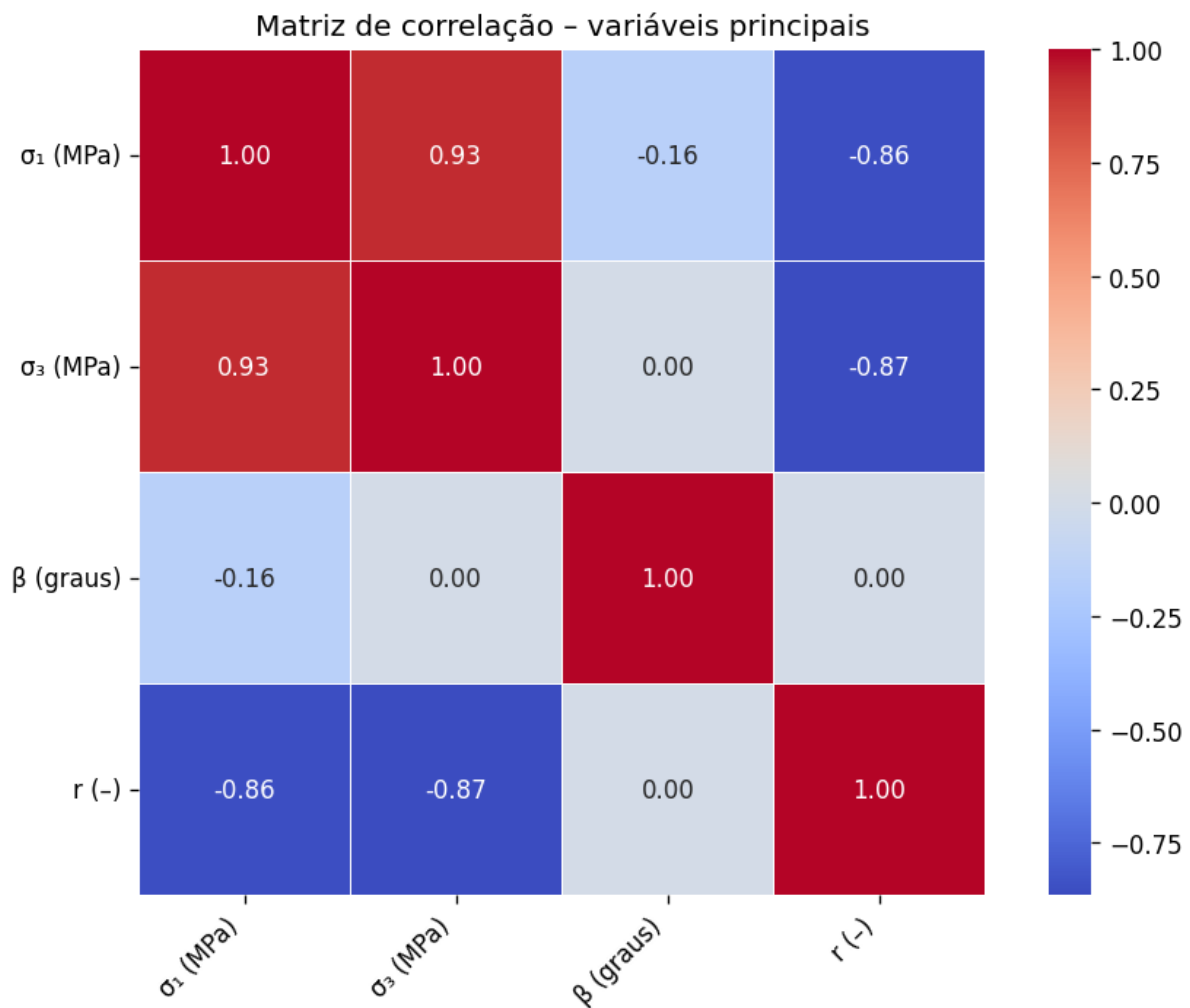
- *Quartzite:*

Estadística descritiva global

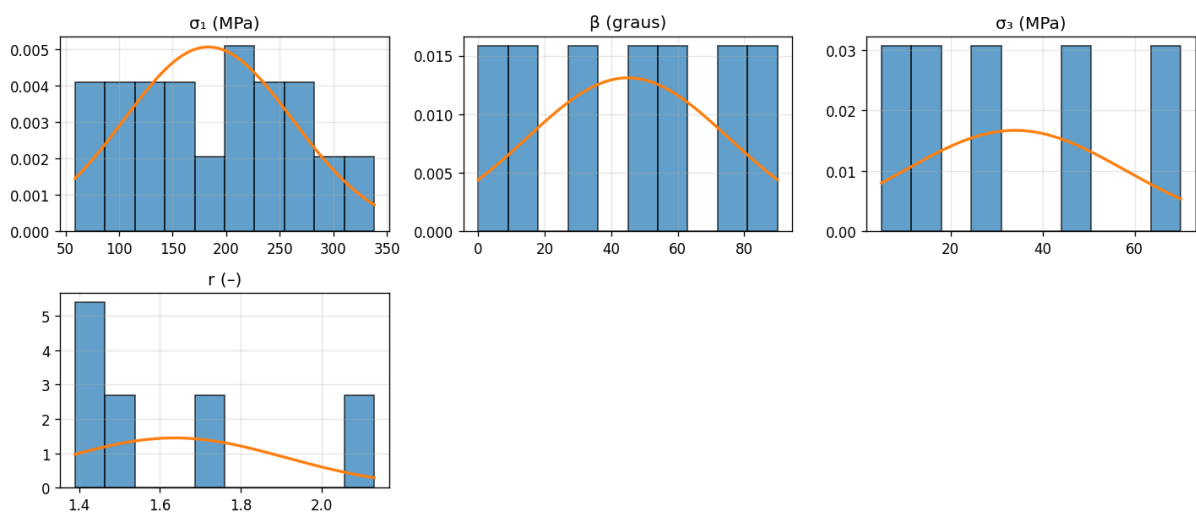
	N	Média	Mediana	Moda	Desvio-padrão	Variância	CV (%)	Mínimo	Q1	Q3	Máximo	Assimetria	Curtose
σ_1 (MPa)	35.00	183.27	173.50	198.50	78.65	6186.57	42.92	58.80	118.35	245.55	338.20	0.15	-1.04
β_{eff} (graus, periodo 180°)	35.00	45.00	45.00	0.00	30.44	926.47	67.64	0.00	15.00	75.00	90.00	0.00	-1.26
σ_3 (MPa)	35.00	34.00	30.00	5.00	23.88	570.29	70.24	5.00	15.00	50.00	70.00	0.31	-1.33
r (-)	35.00	1.64	1.53	1.39	0.28	0.08	16.82	1.39	1.42	1.71	2.13	0.99	-0.48
σ_c (MPa)	35.00	56.80	56.80	56.80	0.00	0.00	0.00	56.80	56.80	56.80	56.80	0.00	0.00
Coesão (MPa)	35.00	22.60	22.60	22.60	0.00	0.00	0.00	22.60	22.60	22.60	22.60	0.00	0.00
Ângulo de atrito (°)	35.00	30.60	30.60	30.60	0.00	0.00	0.00	30.60	30.60	30.60	30.60	0.00	0.00
$\sigma_1 - \sigma_3$ (MPa)	35.00	149.27	148.50	120.00	57.08	3258.31	38.24	53.80	103.35	186.75	268.20	0.14	-0.84
σ_1/σ_3 (-)	35.00	8.03	5.74	3.99	5.15	26.50	64.13	3.47	4.54	9.61	25.00	1.76	2.89

Estatísticas por nível de σ_1 - Variáveis de resistência

Variável	N	σ_1 (MPa)						$\sigma_1 - \sigma_3$ (MPa)						σ_1/σ_3 (-)					
Estatística		Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo
σ_3 (MPa)																			
5	7	83.39	77.90	23.70	561.72	58.80	125.00	78.39	72.90	23.70	561.72	53.80	120.00	16.68	15.58	4.74	22.47	11.76	25.00
15	7	129.19	119.10	26.03	677.35	101.50	173.50	114.19	104.10	26.03	677.35	86.50	158.50	8.61	7.94	1.74	3.01	6.77	11.57
30	7	180.67	172.10	29.02	842.02	147.10	225.00	150.67	142.10	29.02	842.02	117.10	195.00	6.02	5.74	0.97	0.94	4.90	7.50
50	7	236.99	227.90	32.16	1034.43	198.50	282.40	186.99	177.90	32.16	1034.43	148.50	232.40	4.74	4.56	0.64	0.41	3.97	5.65
70	7	286.11	279.40	34.29	1175.49	242.60	338.20	216.11	209.40	34.29	1175.49	172.60	268.20	4.09	3.99	0.49	0.24	3.47	4.83



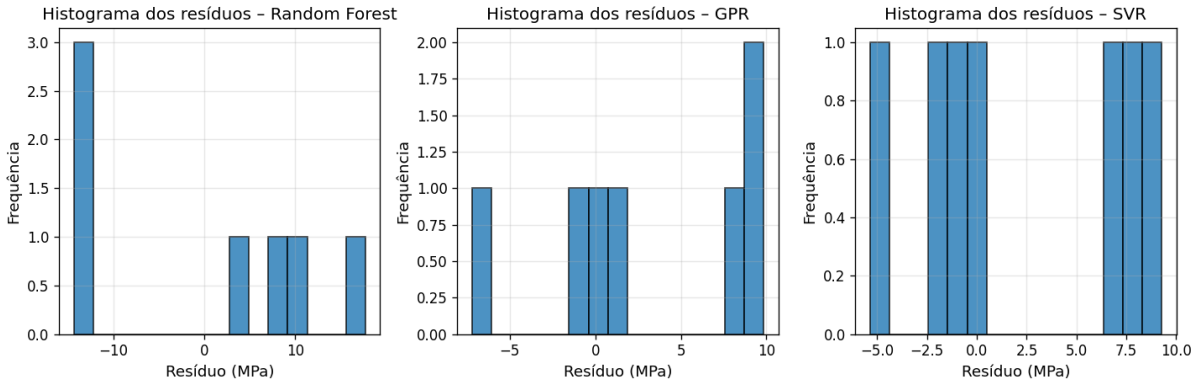
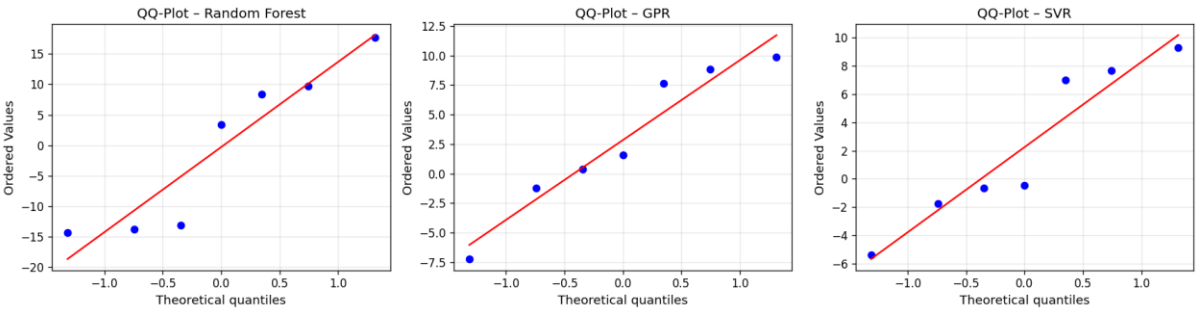
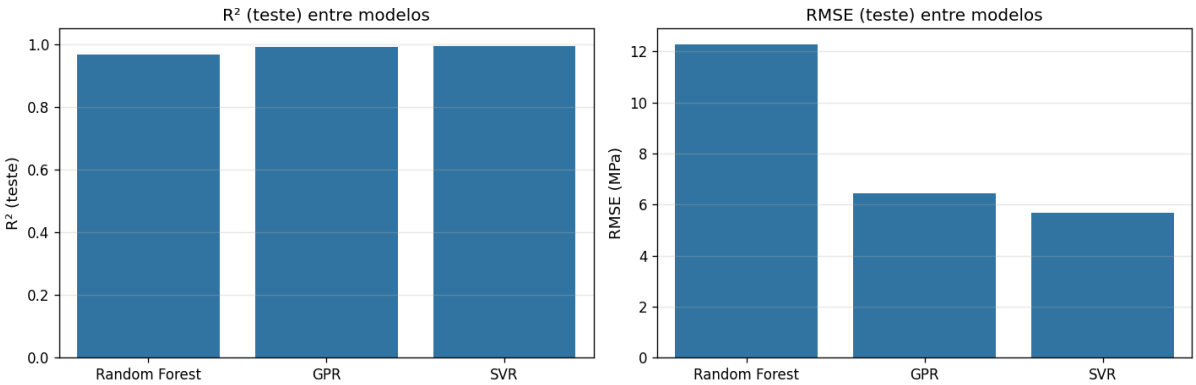
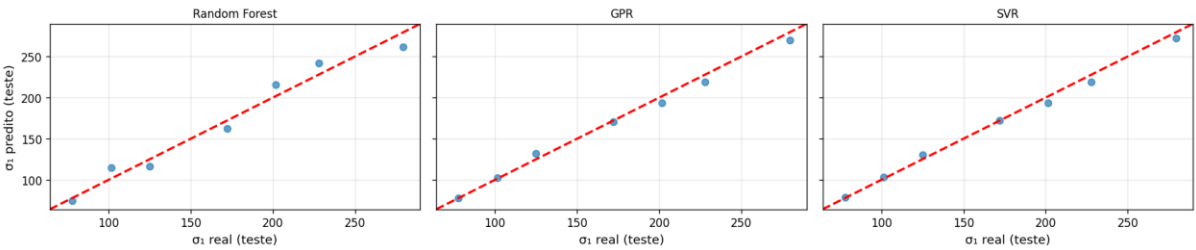
Histogramas + curva Normal - variáveis base

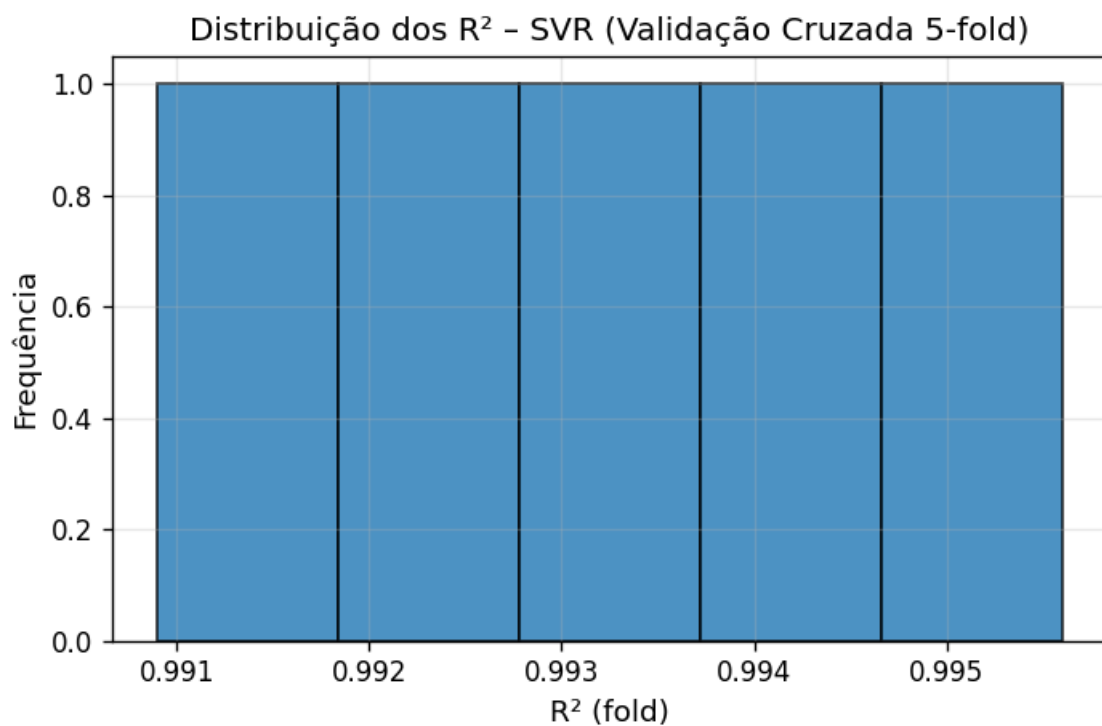
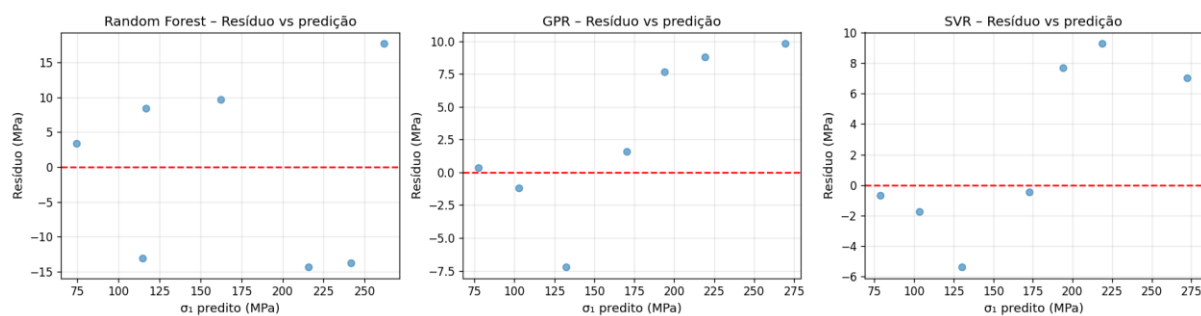


Métricas globais dos modelos (incluindo validação cruzada)

	Modelo	R2_treino	R2_teste	RMSE_treino	RMSE_teste	MAE_treino	MAE_teste	R2_CV_medio	R2_CV_std
0	Random Forest	0.99	0.97	7.50	12.30	6.15	11.50	0.87	0.13
1	GPR	1.00	0.99	1.00	6.43	0.79	5.25	1.00	0.00
2	SVR	1.00	0.99	1.35	5.69	0.59	4.60	0.96	0.05

Comparação visual – σ_1 real vs predito (teste)





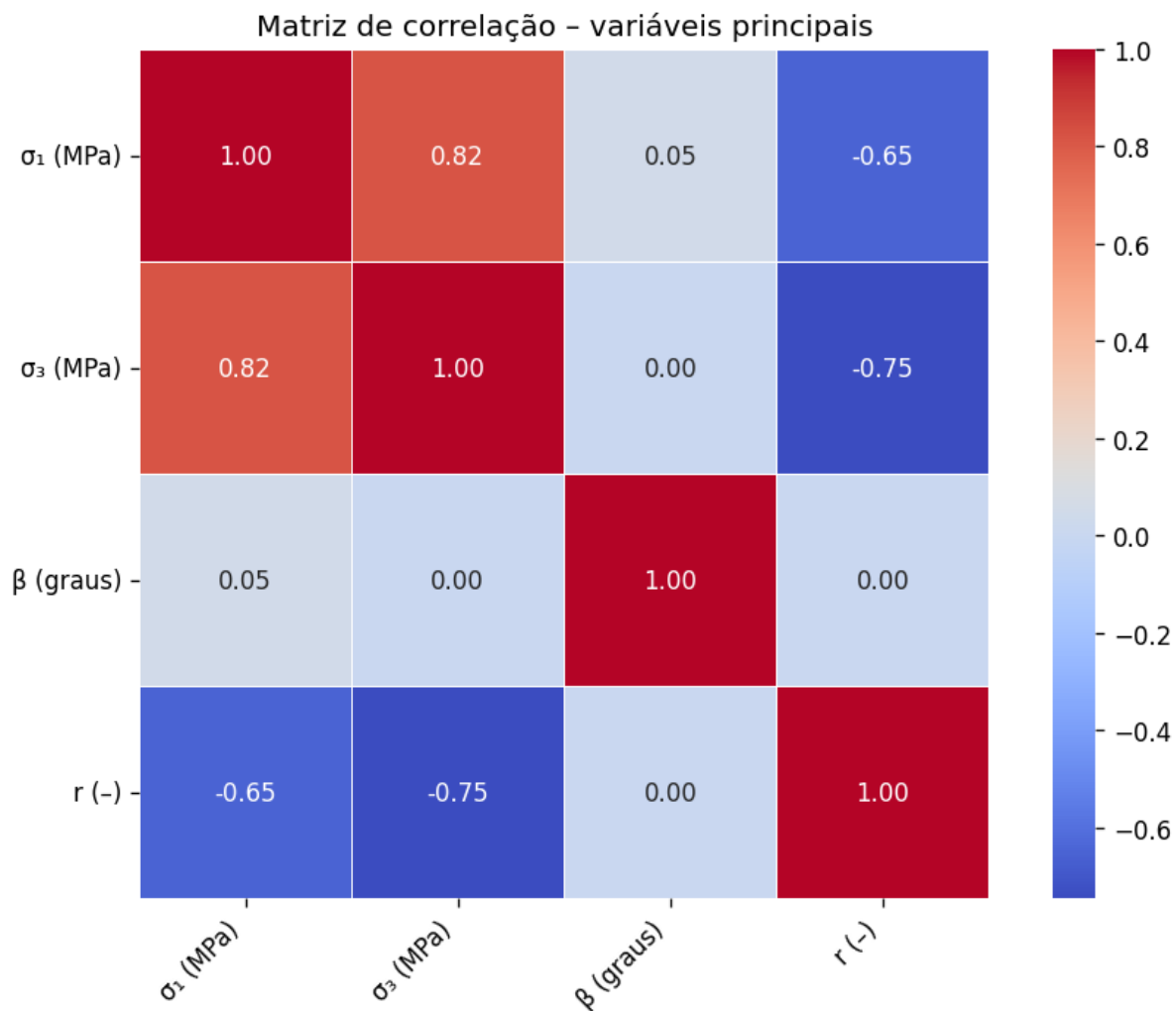
- *Micaceous:*

Estatística descritiva global

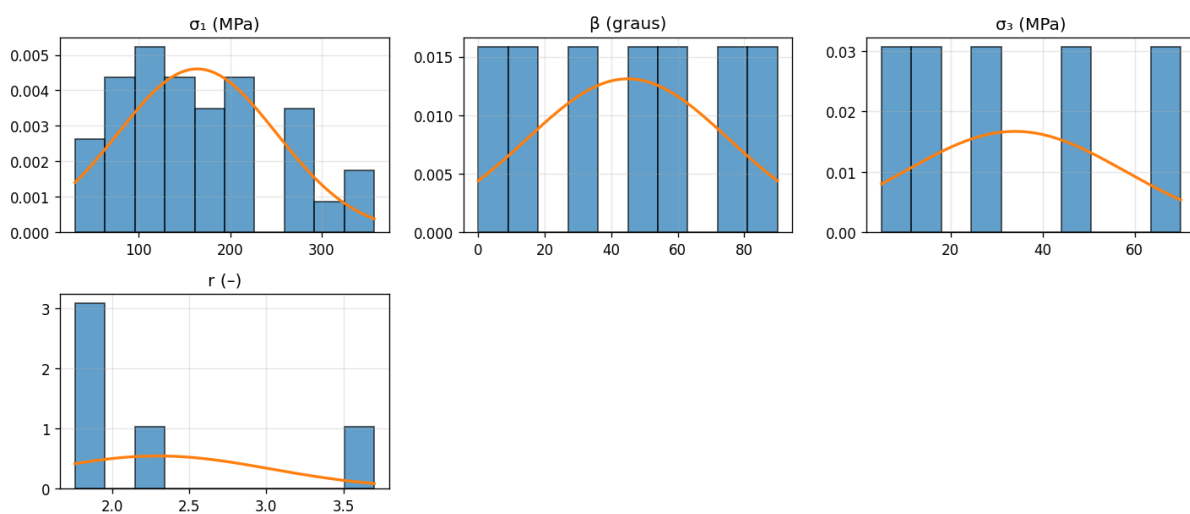
	N	Média	Mediana	Moda	Desvio-padrão	Variância	CV (%)	Mínimo	Q1	Q3	Máximo	Assimetria	Curtose
σ_1 (MPa)	35.00	163.99	149.30	70.50	86.64	7506.11	52.83	30.30	108.00	220.85	357.60	0.53	-0.34
β_{eff} (graus, período 180°)	35.00	45.00	45.00	0.00	30.44	926.47	67.64	0.00	15.00	75.00	90.00	0.00	-1.26
σ_2 (MPa)	35.00	34.00	30.00	5.00	23.88	570.29	70.24	5.00	15.00	50.00	70.00	0.31	-1.33
r (-)	35.00	2.30	1.92	1.76	0.73	0.53	31.82	1.76	1.86	2.24	3.70	1.39	0.18
σ_c (MPa)	35.00	38.80	38.80	38.80	0.00	0.00	0.00	38.80	38.80	38.80	38.80	0.00	0.00
Coesão (MPa)	35.00	18.20	18.20	18.20	0.00	0.00	0.00	18.20	18.20	18.20	18.20	0.00	0.00
Ângulo de atrito (°)	35.00	29.80	29.80	29.80	0.00	0.00	0.00	29.80	29.80	29.80	29.80	0.00	0.00
$\sigma_1 - \sigma_2$ (MPa)	35.00	129.99	111.70	192.90	68.44	4684.63	52.65	25.30	78.00	168.00	307.60	0.72	0.24
σ_1/σ_2 (-)	35.00	6.94	4.98	3.76	4.99	24.88	71.92	2.55	3.80	7.50	23.34	2.15	4.60

Estatísticas por nível de σ_1 – Variáveis de resistência

Variável	N	σ_1 (MPa)						$\sigma_1 - \sigma_2$ (MPa)						σ_1/σ_2 (-)					
Estatística		Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo
σ_2 (MPa)																			
5	7	68.96	70.50	35.38	1251.62	30.30	116.70	63.96	65.50	35.38	1251.62	25.30	111.70	13.79	14.10	7.08	50.06	6.06	23.34
15	7	113.77	112.90	40.87	1670.52	69.70	175.80	98.77	97.90	40.87	1670.52	54.70	160.80	7.58	7.53	2.72	7.42	4.65	11.72
30	7	155.56	144.70	44.12	1946.97	106.10	224.30	125.56	114.70	44.12	1946.97	76.10	194.30	5.19	4.82	1.47	2.16	3.54	7.48
50	7	215.94	203.80	75.90	5760.23	137.90	357.60	165.94	153.80	75.90	5760.23	87.90	307.60	4.32	4.08	1.52	2.30	2.76	7.15
70	7	265.71	262.90	55.88	3122.21	178.80	348.50	195.71	192.90	55.88	3122.21	108.80	278.50	3.80	3.76	0.80	0.64	2.55	4.98



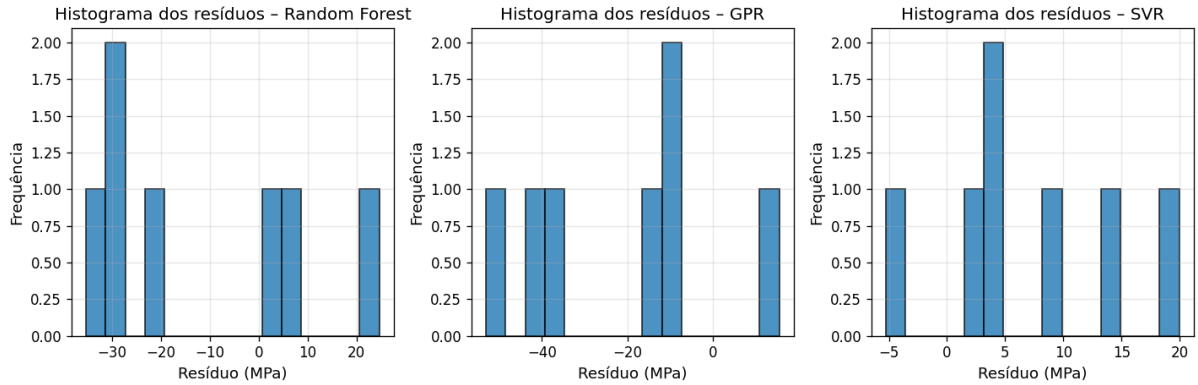
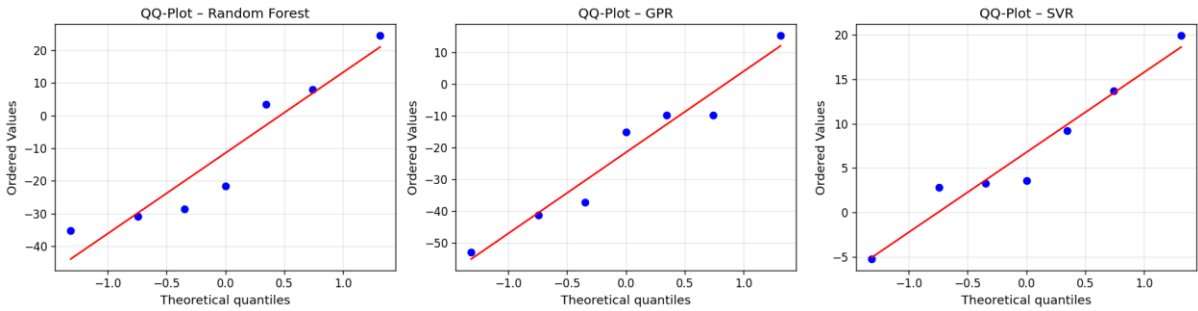
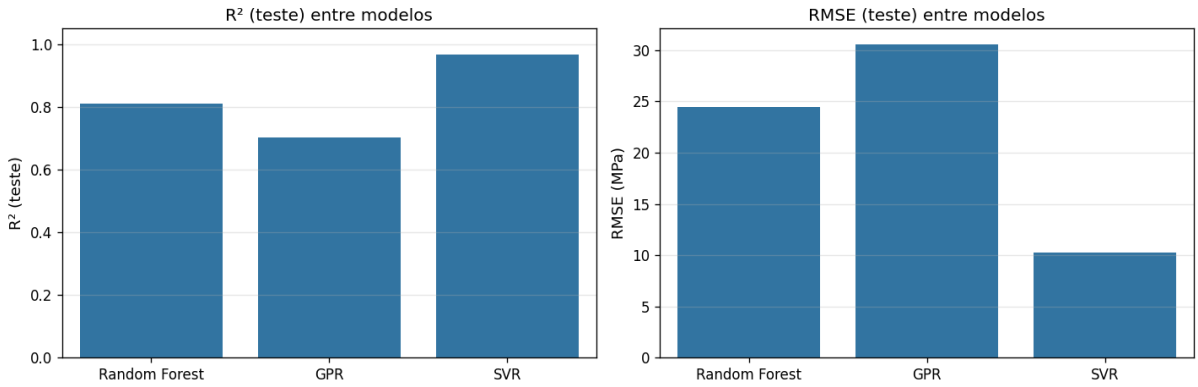
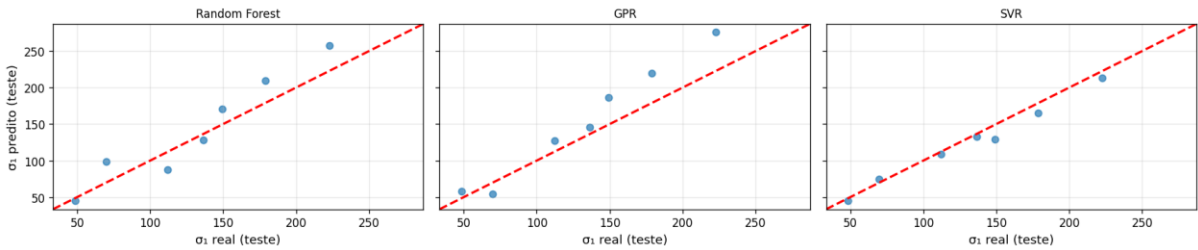
Histogramas + curva Normal - variáveis base

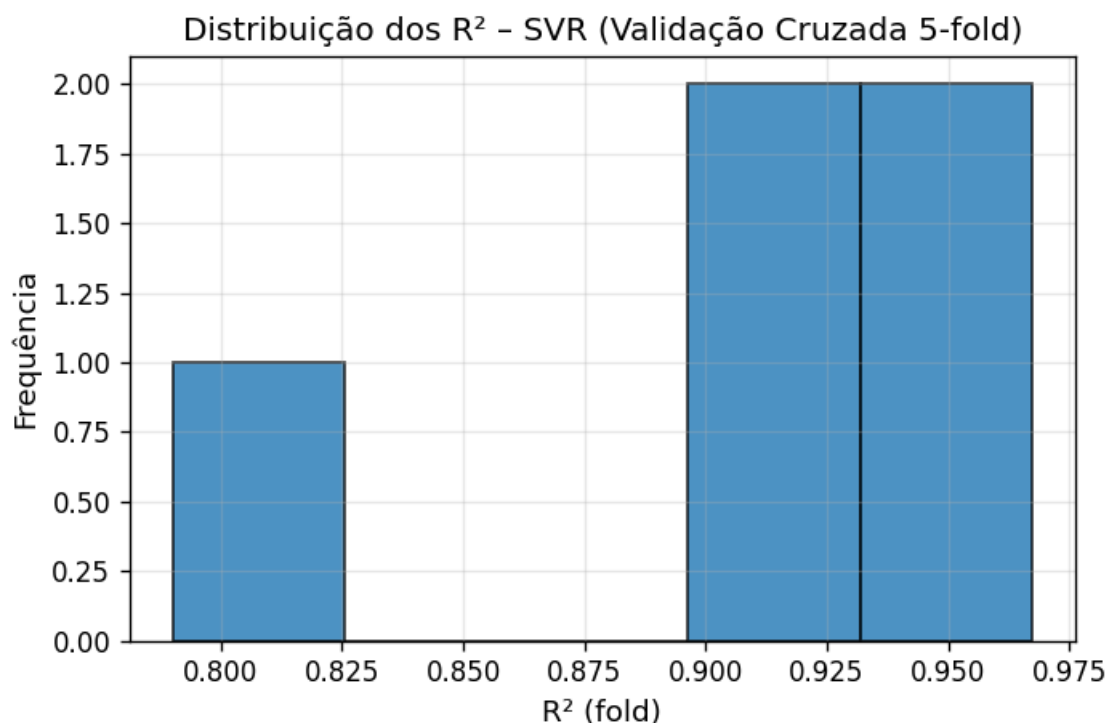
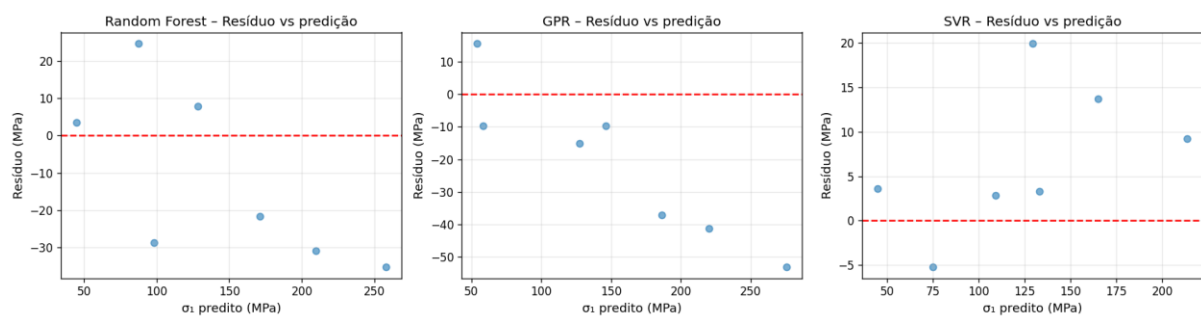


Métricas globais dos modelos (incluindo validação cruzada)

	Modelo	R2_treino	R2_teste	RMSE_treino	RMSE_teste	MAE_treino	MAE_teste	R2_CV_medio	R2_CV_std
0	Random Forest	0.97	0.81	16.30	24.41	12.27	21.79	0.28	0.74
1	GPR	0.98	0.70	11.51	30.60	9.34	25.95	0.54	0.47
2	SVR	0.95	0.97	19.51	10.22	7.44	8.27	0.57	0.62

Comparação visual – σ_1 real vs predito (teste)





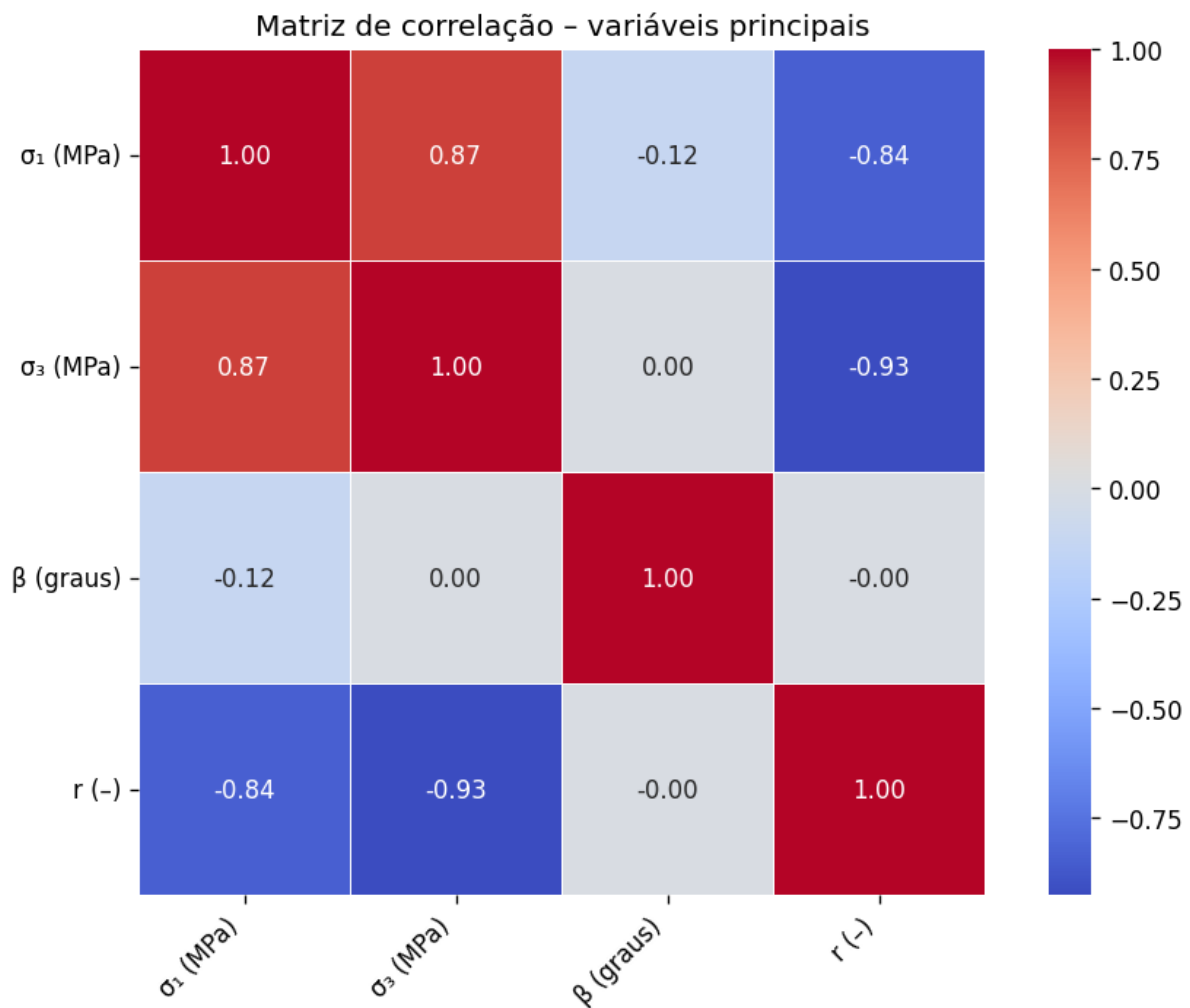
- *Dry:*

Estadística descritiva global

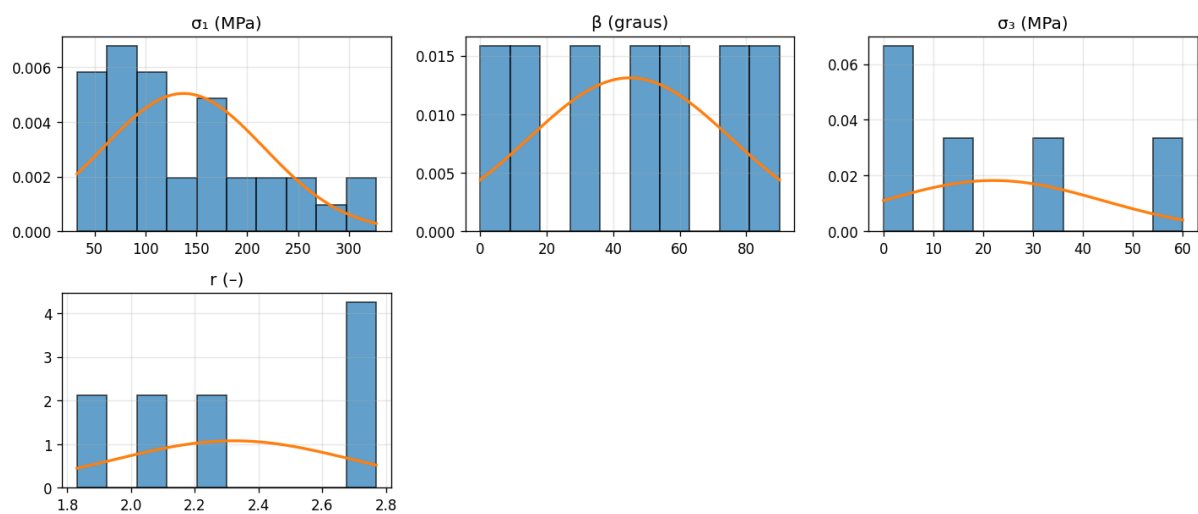
	N	Média	Mediana	Moda	Desvio-padrão	Variância	CV (%)	Mínimo	Q1	Q3	Máximo	Assimetria	Curtose
σ_1 (MPa)	35.00	137.48	118.20	32.70	79.10	6257.45	57.54	32.70	76.25	183.95	326.80	0.78	-0.14
β_{eff} (graus, período 180°)	35.00	45.00	45.00	0.00	30.44	926.47	67.64	0.00	15.00	75.00	90.00	0.00	-1.26
σ_2 (MPa)	35.00	22.00	15.00	0.00	21.90	479.71	99.56	0.00	5.00	30.00	60.00	0.82	-0.71
r (-)	35.00	2.32	2.27	1.83	0.37	0.14	15.97	1.83	2.04	2.70	2.77	0.01	-1.65
σ_c (MPa)	35.00	58.80	58.80	58.80	0.00	0.00	0.00	58.80	58.80	58.80	58.80	0.00	0.00
Coesão (MPa)	35.00	19.30	19.30	19.30	0.00	0.00	0.00	19.30	19.30	19.30	19.30	0.00	0.00
Ângulo de atrito (°)	35.00	31.10	31.10	31.10	0.00	0.00	0.00	31.10	31.10	31.10	31.10	0.00	0.00
$\sigma_1 - \sigma_2$ (MPa)	35.00	115.48	98.80	32.70	61.07	3728.97	52.88	32.70	71.00	152.05	266.80	0.83	0.11
σ_1/σ_2 (-)	28.00	8.53	6.80	2.98	5.57	31.05	65.35	2.98	4.95	9.98	23.64	1.52	1.64

Estatísticas por nível de σ_1 - Variáveis de resistência

Variável	N	σ_1 (MPa)						$\sigma_1 - \sigma_2$ (MPa)						σ_1/σ_2 (-)					
		Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo
σ_2 (MPa)																			
0	7	60.94	58.90	21.07	444.11	32.70	88.00	60.94	58.90	21.07	444.11	32.70	88.00	nan	nan	nan	nan	nan	nan
5	7	79.86	76.80	29.40	864.39	42.70	118.20	74.86	71.80	29.40	864.39	37.70	113.20	15.97	15.36	5.88	34.58	8.54	23.64
15	7	121.81	113.80	37.75	1425.22	75.70	171.90	106.81	98.80	37.75	1425.22	60.70	156.90	8.12	7.59	2.52	6.33	5.05	11.46
30	7	176.16	165.80	46.17	2131.93	120.50	246.30	146.16	135.80	46.17	2131.93	90.50	216.30	5.87	5.53	1.54	2.37	4.02	8.21
60	7	248.61	239.60	58.69	3444.30	178.60	326.80	188.61	179.60	58.69	3444.30	118.60	266.80	4.14	3.99	0.98	0.96	2.98	5.45



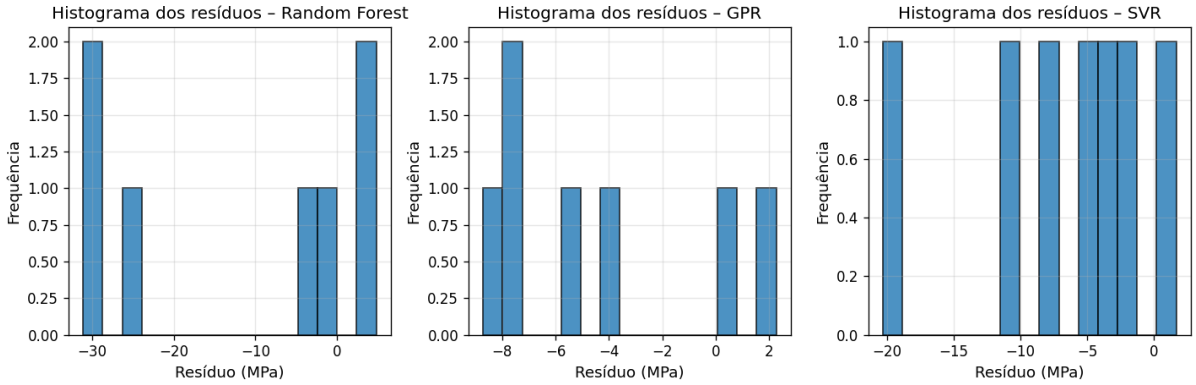
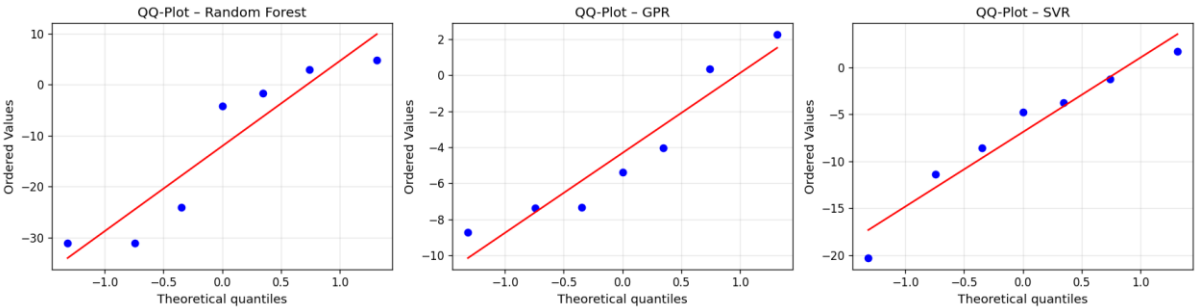
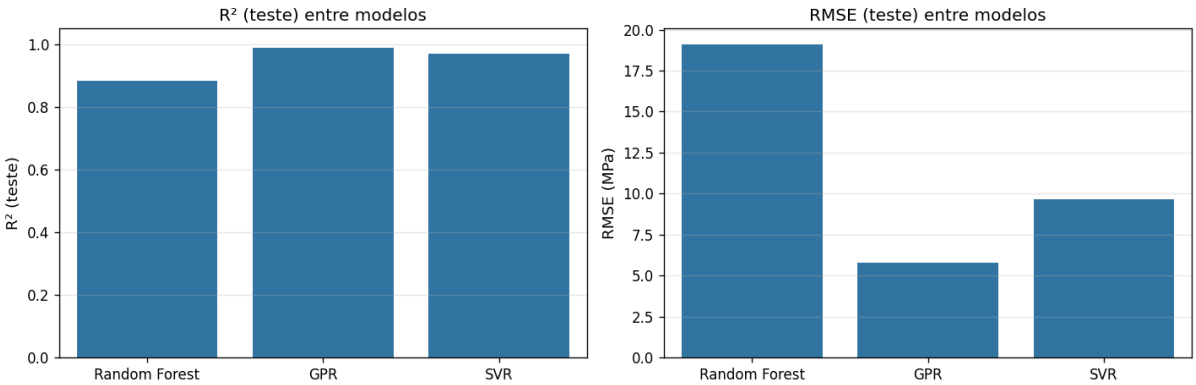
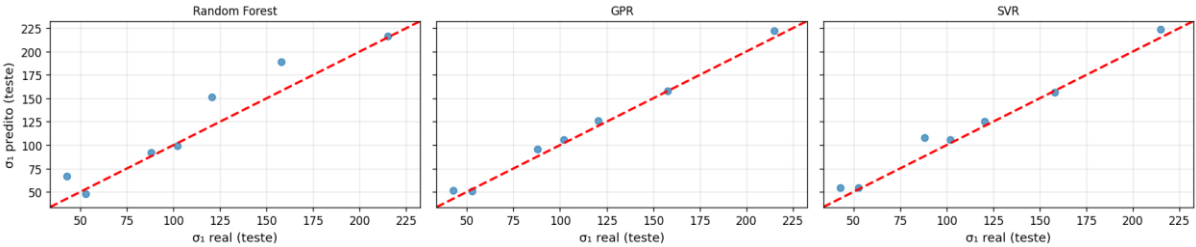
Histogramas + curva Normal - variáveis base

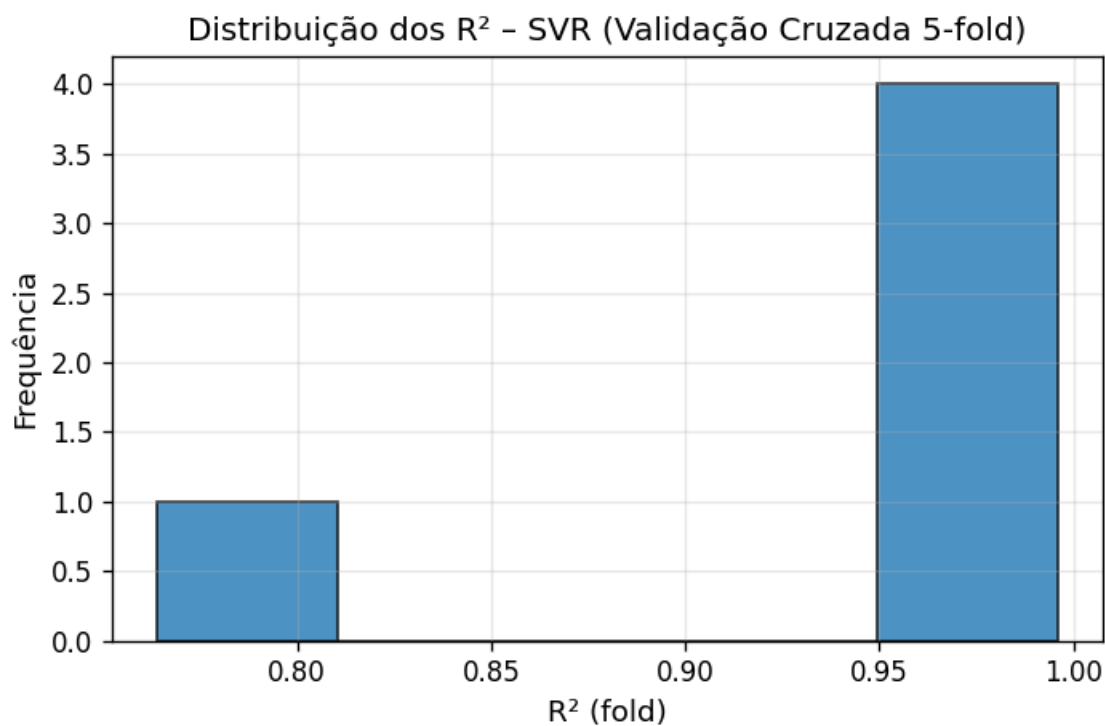
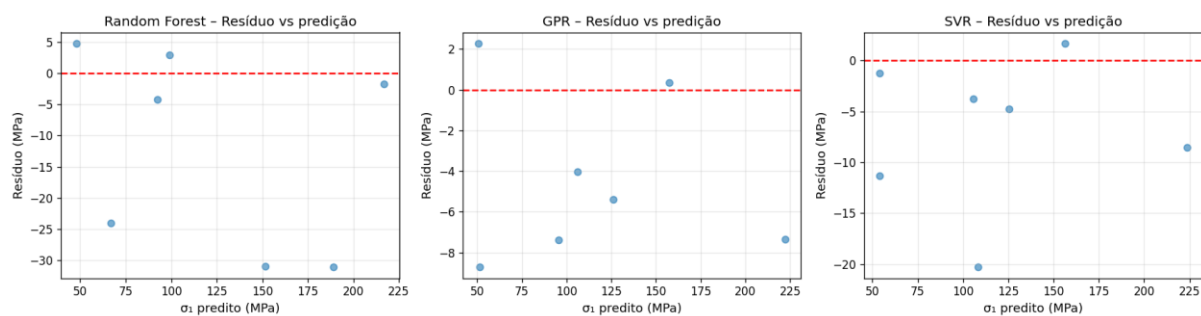


Métricas globais dos modelos (incluindo validação cruzada)

	Modelo	R2_treino	R2_teste	RMSE_treino	RMSE_teste	MAE_treino	MAE_teste	R2_CV_medio	R2_CV_std
0	Random Forest	0.98	0.88	11.97	19.12	8.86	14.26	0.66	0.26
1	GPR	1.00	0.99	4.08	5.79	3.18	5.07	0.97	0.02
2	SVR	0.99	0.97	6.17	9.68	3.36	7.38	0.91	0.12

Comparação visual - σ_1 real vs predito (teste)





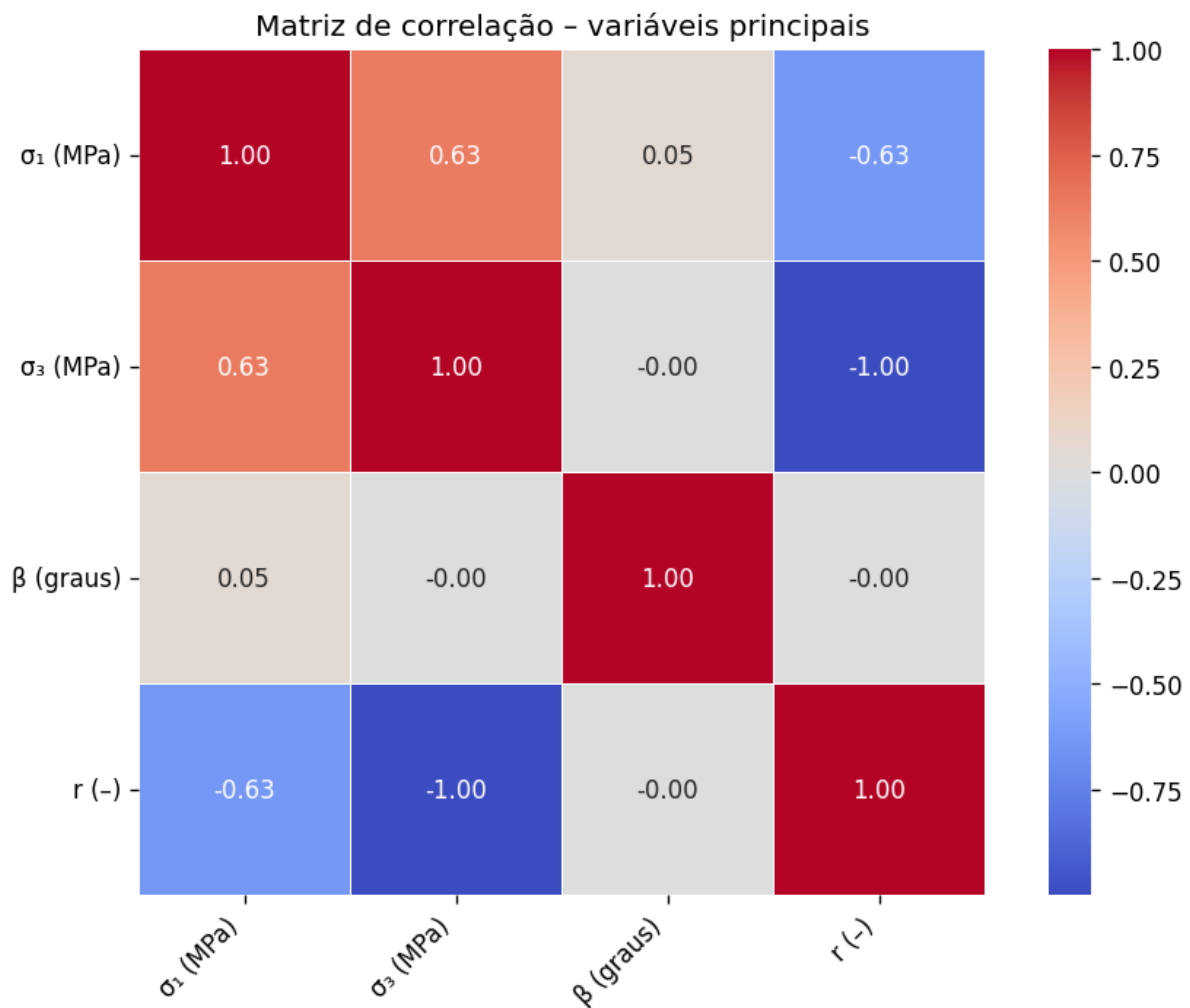
- *Moretown:*

Estadística descritiva global

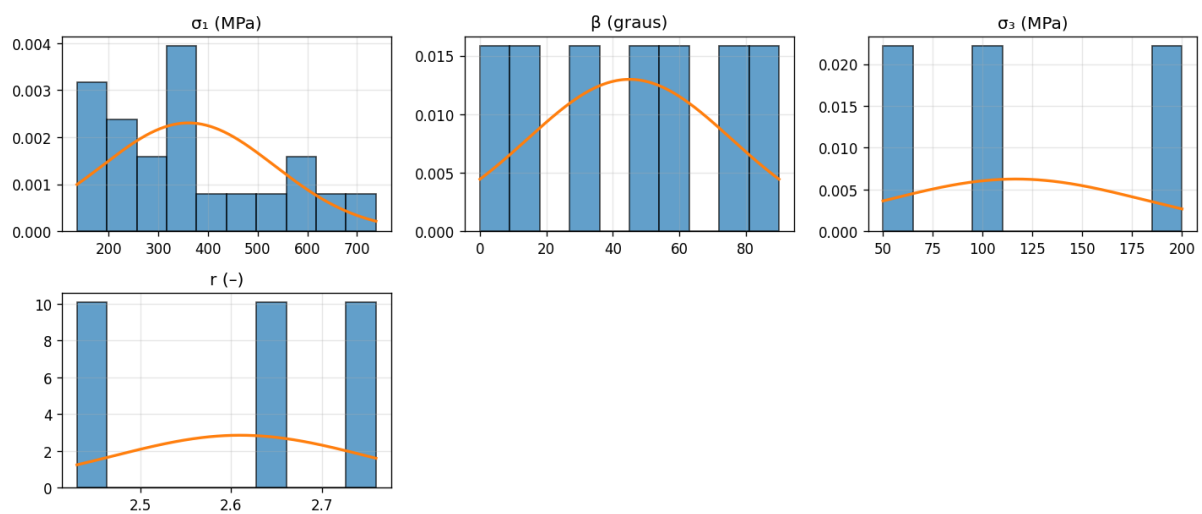
	N	Média	Mediana	Moda	Desvio-padrão	Variância	CV (%)	Mínimo	Q1	Q3	Máximo	Assimetria	Curtose
σ_1 (MPa)	21.00	360.45	346.80	135.90	173.12	29970.07	48.03	135.90	223.50	472.00	738.30	0.69	-0.38
β_{eff} (graus, período 180°)	21.00	45.00	45.00	0.00	30.74	945.00	68.31	0.00	15.00	75.00	90.00	0.00	-1.26
σ_3 (MPa)	21.00	116.67	100.00	50.00	63.90	4083.33	54.77	50.00	50.00	200.00	200.00	0.41	-1.58
r (-)	21.00	2.61	2.64	2.43	0.14	0.02	5.35	2.43	2.43	2.76	2.76	-0.34	-1.58
σ_c (MPa)	21.00	140.90	140.90	140.90	0.00	0.00	0.00	140.90	140.90	140.90	140.90	0.00	0.00
Coesão (MPa)	21.00	61.10	61.10	61.10	0.00	0.00	0.00	61.10	61.10	61.10	61.10	0.00	0.00
Ângulo de atrito (°)	21.00	15.30	15.30	15.30	0.00	0.00	0.00	15.30	15.30	15.30	15.30	0.00	0.00
$\sigma_1 - \sigma_3$ (MPa)	21.00	243.78	219.90	68.20	141.53	20032.07	58.06	68.20	124.30	357.80	538.30	0.49	-0.98
σ_1/σ_3 (-)	21.00	3.59	3.26	1.34	1.77	3.14	49.38	1.34	2.32	4.42	7.97	1.20	1.23

Estadísticas por nível de σ_1 - Variáveis de resistência

Variável	N	σ_1 (MPa)						$\sigma_1 - \sigma_3$ (MPa)						σ_1/σ_3 (-)					
Estatística		Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo	Média	Mediana	Desvio-padrão	Variância	Mínimo	Máximo
σ_3 (MPa)																			
50	7	246.71	220.80	105.96	11227.20	135.90	398.70	196.71	170.80	105.96	11227.20	85.90	348.70	4.93	4.42	2.12	4.49	2.72	7.97
100	7	330.74	346.80	127.22	16183.70	178.80	509.50	230.74	246.80	127.22	16183.70	78.80	409.50	3.31	3.47	1.27	1.62	1.79	5.10
200	7	503.89	557.80	179.90	32365.16	268.20	738.30	303.89	357.80	179.90	32365.16	68.20	538.30	2.52	2.79	0.90	0.81	1.34	3.69



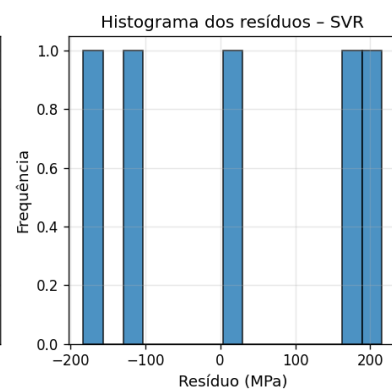
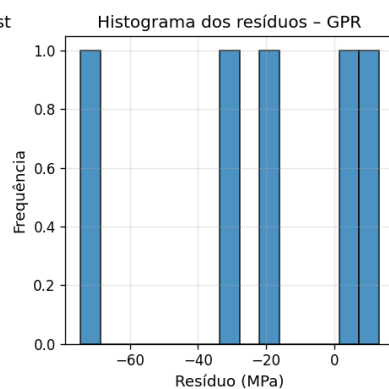
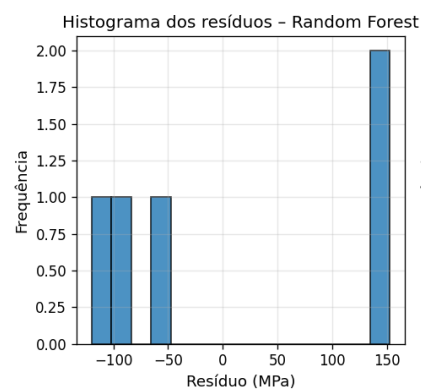
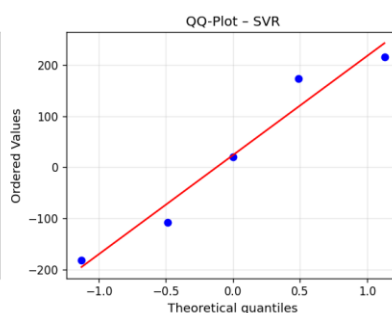
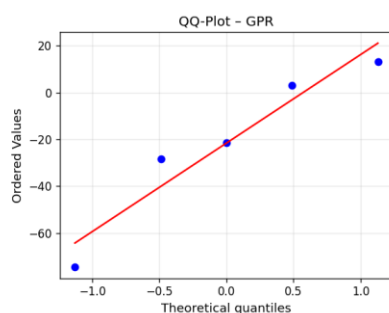
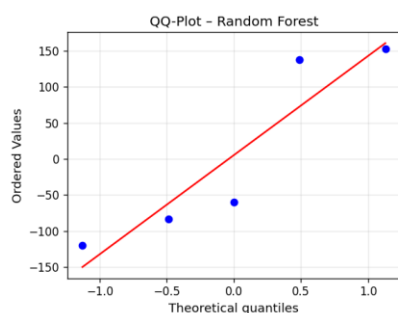
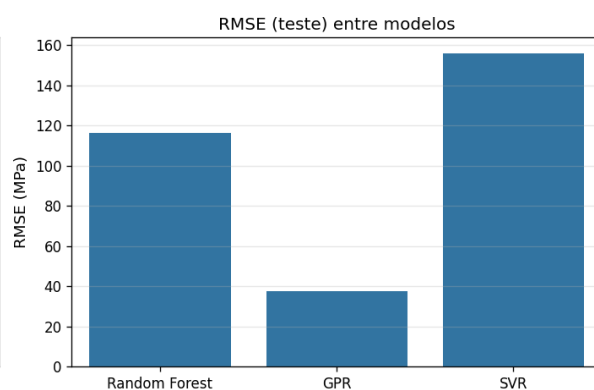
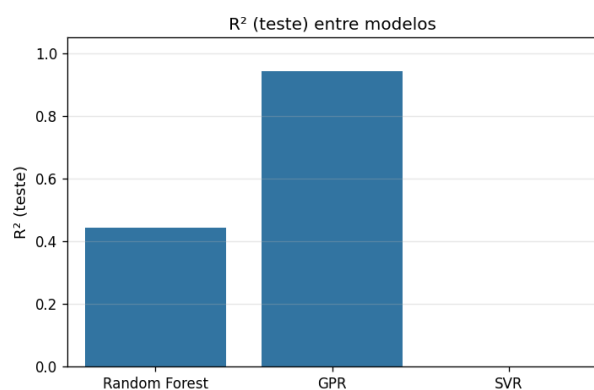
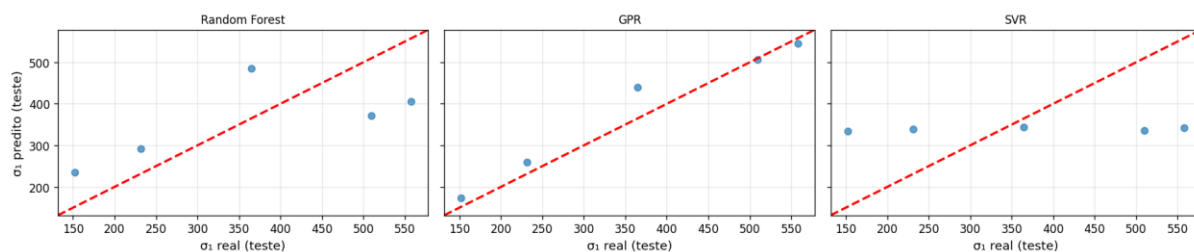
Histogramas + curva Normal - variáveis base

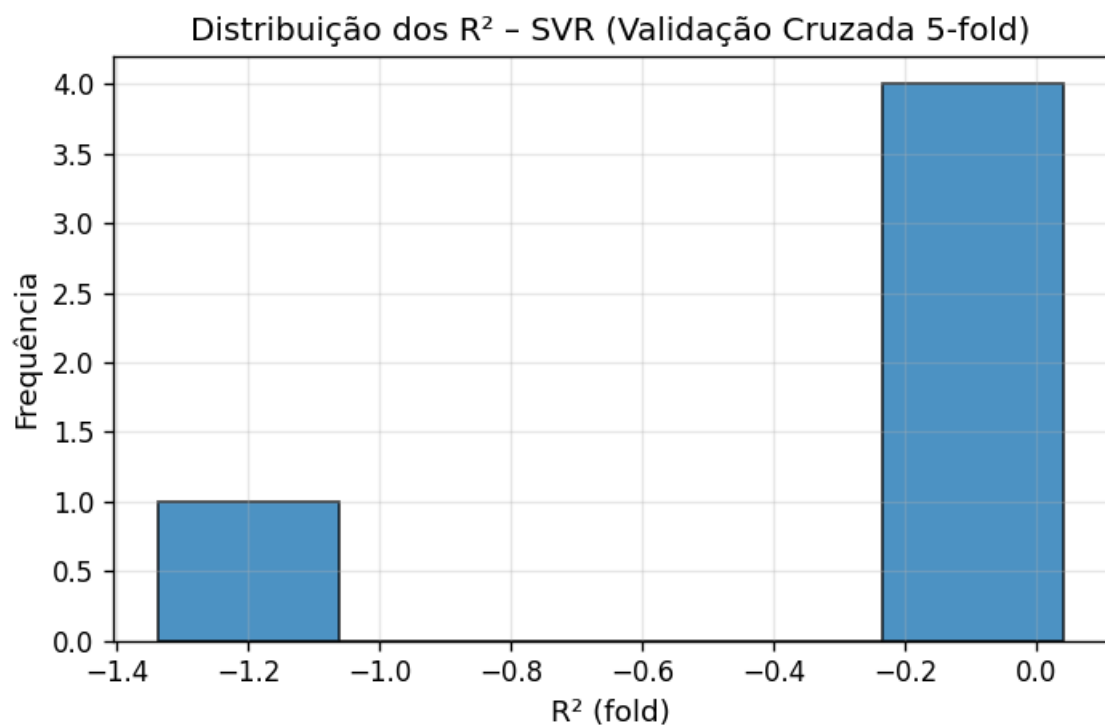
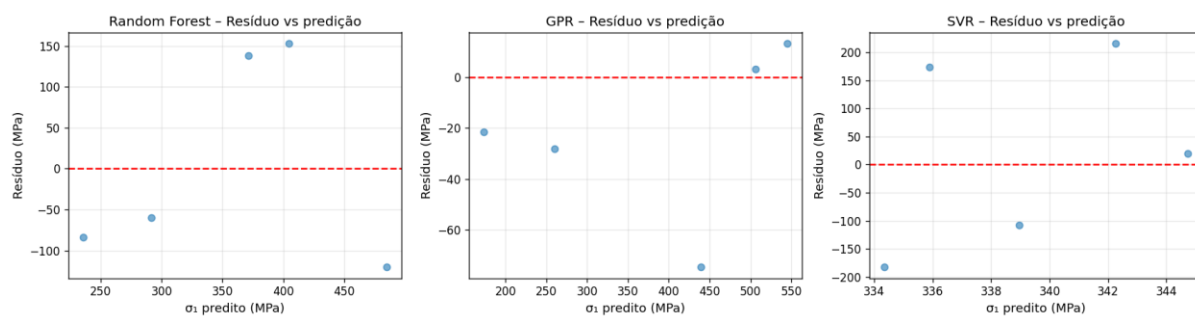


Métricas globais dos modelos (incluindo validação cruzada)

	Modelo	R2_treino	R2_teste	RMSE_treino	RMSE_teste	MAE_treino	MAE_teste	R2_CV_medio	R2_CV_std
0	Random Forest	0.93	0.44	46.30	116.18	36.20	110.98	-1.12	1.65
1	GPR	1.00	0.94	0.02	37.40	0.01	28.07	0.28	0.62
2	SVR	0.01	-0.00	171.68	156.08	132.58	139.79	-3.51	4.06

Comparação visual - σ_1 real vs predito (teste)





2.6. Bancos de dados:

Rocha	Lito	Ref	Ano	Sigma1	Beta	Sigma3	r	Sigmac	Coesao	Atrito
1	Carbonaceous Phyllite	Singh	1988	330	90	70	1,31	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	263,1	90	50	1,37	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	203,1	90	30	1,48	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	144,6	90	15	1,71	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	98,5	90	5	1,97	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	286,7	75	70	1,31	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	231,3	75	50	1,37	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	171,3	75	30	1,48	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	119	75	15	1,71	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	74,4	75	5	1,97	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	262,6	60	70	1,31	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	204,1	60	50	1,37	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	148,7	60	30	1,48	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	98	60	15	1,71	56,5	22	31,1

1	Carbonaceous Phyllite	Singh	1988	62,6	60	5	1,97	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	255,4	45	70	1,31	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	201,5	45	50	1,37	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	149,2	45	30	1,48	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	100	45	15	1,71	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	63,1	45	5	1,97	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	275,9	30	70	1,31	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	223,6	30	50	1,37	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	169,8	30	30	1,48	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	119	30	15	1,71	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	82,1	30	5	1,97	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	308,7	15	70	1,31	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	251,8	15	50	1,37	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	196,4	15	30	1,48	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	147,2	15	15	1,71	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	101,1	15	5	1,97	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	335,4	0	70	1,31	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	276,9	0	50	1,37	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	220	0	30	1,48	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	167,7	0	15	1,71	56,5	22	31,1
1	Carbonaceous Phyllite	Singh	1988	123,1	0	5	1,97	56,5	22	31,1

Rocha	Lito	Ref	Ano	Sigma1	Beta	Sigma3	r	Sigmac	Coesao	Atrito
2	Quartzite Phyllite	Singh	1988	302,9	90	70	1,39	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	266,2	90	50	1,42	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	198,5	90	30	1,53	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	139,7	90	15	1,71	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	86,8	90	5	2,13	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	279,4	75	70	1,39	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	226,5	75	50	1,42	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	169,1	75	30	1,53	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	117,6	75	15	1,71	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	69,1	75	5	2,13	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	248,5	60	70	1,39	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	198,5	60	50	1,42	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	147,1	60	30	1,53	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	101,5	60	15	1,71	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	58,8	60	5	2,13	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	242,6	45	70	1,39	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	201,5	45	50	1,42	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	150	45	30	1,53	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	104,4	45	15	1,71	56,8	22,6	30,6

2	Quartzite Phyllite	Singh	1988	63,2	45	5	2,13	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	279,4	30	70	1,39	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	227,9	30	50	1,42	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	172,1	30	30	1,53	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	119,1	30	15	1,71	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	77,9	30	5	2,13	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	311,8	15	70	1,39	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	255,9	15	50	1,42	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	202,9	15	30	1,53	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	148,5	15	15	1,71	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	102,9	15	5	2,13	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	338,2	0	70	1,39	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	282,4	0	50	1,42	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	225	0	30	1,53	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	173,5	0	15	1,71	56,8	22,6	30,6
2	Quartzite Phyllite	Singh	1988	125	0	5	2,13	56,8	22,6	30,6

Rocha	Lito	Ref	Ano	Sigma1	Beta	Sigma3	r	Sigmat	Coesao	Atrito
3	Micaceous Phyllite	Singh	1988	348,5	90	70	1,76	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	357,6	90	50	1,92	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	224,3	90	30	1,86	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	175,8	90	15	2,24	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	116,7	90	5	3,7	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	269	75	70	1,76	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	203,8	75	50	1,92	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	144,7	75	30	1,86	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	112,9	75	15	2,24	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	70,5	75	5	3,7	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	178,8	60	70	1,76	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	137,9	60	50	1,92	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	106,1	60	30	1,86	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	69,7	60	15	2,24	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	30,3	60	5	3,7	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	262,9	45	70	1,76	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	149,3	45	50	1,92	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	109,9	45	30	1,86	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	70,5	45	15	2,24	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	31,1	45	5	3,7	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	222,7	30	70	1,76	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	178,8	30	50	1,92	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	136,4	30	30	1,86	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	90,9	30	15	2,24	38,8	18,2	29,8

3	Micaceous Phyllite	Singh	1988	48,5	30	5	3,7	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	262,9	15	70	1,76	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	219	15	50	1,92	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	170,5	15	30	1,86	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	120,5	15	15	2,24	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	73,5	15	5	3,7	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	315,2	0	70	1,76	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	265,2	0	50	1,92	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	197	0	30	1,86	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	156,1	0	15	2,24	38,8	18,2	29,8
3	Micaceous Phyllite	Singh	1988	112,1	0	5	3,7	38,8	18,2	29,8

Rocha	Lito	Ref	Ano	Sigma1	Beta	Sigma3	<i>r</i>	Sigmac	Coesao	Atrito
4	Phyllites dry	Kumar	2006	80,2	90	0	2,7	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	107,6	90	5	2,77	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	164,1	90	15	2,27	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	211,1	90	30	2,04	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	311,7	90	60	1,83	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	58,9	75	0	2,7	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	76,8	75	5	2,77	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	113,8	75	15	2,27	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	165,8	75	30	2,04	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	239,6	75	60	1,83	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	32,7	60	0	2,7	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	42,7	60	5	2,77	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	75,7	60	15	2,27	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	127,7	60	30	2,04	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	178,6	60	60	1,83	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	38,8	45	0	2,7	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	48,3	45	5	2,77	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	85,2	45	15	2,27	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	120,5	45	30	2,04	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	189,3	45	60	1,83	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	52,8	30	0	2,7	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	66,2	30	5	2,77	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	102	30	15	2,27	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	157,9	30	30	2,04	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	215	30	60	1,83	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	75,2	15	0	2,7	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	99,2	15	5	2,77	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	140	15	15	2,27	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	203,8	15	30	2,04	58,8	19,3	31,1

4	Phyllites dry	Kumar	2006	279,3	15	60	1,83	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	88	0	0	2,7	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	118,2	0	5	2,77	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	171,9	0	15	2,27	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	246,3	0	30	2,04	58,8	19,3	31,1
4	Phyllites dry	Kumar	2006	326,8	0	60	1,83	58,8	19,3	31,1

Rocha	Lito	Ref	Ano	Sigma1	Beta	Sigma3	r	Sigmatc	Coesao	Atrito
5	Moretown Phyllite	Donath	1972	375,4	0	50	2,76	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	472	0	100	2,64	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	652,5	0	200	2,43	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	220,8	15	50	2,76	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	346,8	15	100	2,64	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	609,6	15	200	2,43	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	174,3	30	50	2,76	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	231,5	30	100	2,64	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	364,7	30	200	2,43	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	135,9	45	50	2,76	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	178,8	45	100	2,64	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	268,2	45	200	2,43	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	152	60	50	2,76	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	223,5	60	100	2,64	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	336,1	60	200	2,43	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	269,9	75	50	2,76	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	353,1	75	100	2,64	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	557,8	75	200	2,43	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	398,7	90	50	2,76	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	509,5	90	100	2,64	140,9	61,1	15,3
5	Moretown Phyllite	Donath	1972	738,3	90	200	2,43	140,9	61,1	15,3