

**UMA PROPOSTA DE MÉTODO PREDITIVO PARA PRODUÇÃO ANUAL DE PETRÓLEO NO  
MERCADO BRASILEIRO**

**HARADA FRANCO HAMILTON**  
FUNDAÇÃO INSTITUTO DE ADMINISTRAÇÃO - FIA  
francohamilton@yahoo.com.br

# UMA PROPOSTA DE MÉTODO PREDITIVO PARA PRODUÇÃO ANUAL DE PETRÓLEO NO MERCADO BRASILEIRO

## 1. INTRODUÇÃO

O mercado de Óleo & Gás, mercado de grande relevância na economia mundial, abrange serviços, insumos e equipamentos *onshore* e *offshore*, construção civil e naval, áreas de pesquisa e desenvolvimento e diversas outras atividades necessárias para o progresso do setor no que se refere desde a exploração e refino do petróleo até a comercialização de seus subprodutos.

No Brasil, este mercado orbita em torno da PETROBRAS, empresa de relevante importância nacional e internacional, que gera milhares de empregos diretos e indiretos, movimentando vultosas quantias em investimentos para se atingir o aumento da exploração, produção e refino de petróleo, mesmo com todos os desdobramentos da Operação Lava Jato.

Visando prever como o mercado de Óleo & Gás poderá se comportar em um determinado período, a previsão da produção de petróleo no Brasil pode ser um indicativo da necessidade de investimentos em produtos e serviços que orbitam em torno desse mercado ou até processos de desinvestimentos em futuras épocas de crise, auxiliando o investidor e/ou o empreendedor nas tomadas de decisões.

Desta forma, buscou-se a obtenção de um método preditivo que forneça como dado de saída – output – a produção anual brasileira de petróleo para um determinado ano. Por meio do método estabelecido, será analisada a demanda para o mercado de óleo e gás nacional com a previsão da necessidade de investimentos em refinarias, navios-plataforma, plataformas de exploração e todo o suporte necessário para estas atividades macro, desde fabricantes de microcomponentes eletrônicos até estaleiros de grande porte para atendimento a toda a cadeia de fornecimento deste mercado, além das áreas de engenharia e serviços de construção e montagem, e, das áreas de pesquisa e desenvolvimento. Pode ser também mencionada áreas indiretas, como setores de serviços como rede hoteleira, imobiliária e alimentícia, transportes (terrestres, marítimos e aéreos) e de segurança e limpeza.

Outro ponto relevante é quanto ao atendimento ao conteúdo local nos fornecimentos (proporção dos investimentos nacionais para um determinado bem ou serviço que corresponde à parcela de participação da indústria nacional na produção desse bem ou serviço (PROMINP, 2015)), que se mantido este incentivo, aumenta ainda mais a importância de previsões com alto grau de assertividade, pois estimula a economia nacional com diversos fornecedores brasileiros movimentando o mercado de óleo e gás e todo o mercado suporte a ele.

Vale também lembrar que pode ser ferramenta para identificação de investimentos, como por exemplo, em equipamentos ou serviços específicos a área e que ainda haja demanda após a análise do *market share*, comparações entre modelos de negócios, negociações de financiamentos, entre outros.

Quanto a metodologia empregada, foi escolhido o método preditivo através da regressão linear múltipla, pois dadas as informações de entrada, fazendo as considerações necessárias e o tratamento adequado dos dados, devolverão valores futuros previstos pelo modelo obtido. Estas previsões serão aproximações que permitirão aumentar a acuracidade na tomada de decisões e a velocidade na identificação da melhor solução.

Para tanto foram coletados dados de relevância na área econômica, produtiva e de desenvolvimento social brasileiro (dados macroeconômicos, dados de produção, faturamento e consumo de alguns setores produtivos e dados diversos, como o INCC, financiamentos do BNDES para diversos investimentos, salário mínimo, população brasileira e PEA), dados estes de entrada, que com uma série histórica adequada deu robustez a análise para que com todos os dados obtidos fosse realizada uma regressão linear múltipla destes para se chegar ao modelo desejado.

Por mais que fontes renováveis de energia estejam em estudo, tais como a energia solar, geração de hidrogênio a partir da água, cogeração de energia, entre outras (PLANETA SUSTENTÁVEL, 2015), o petróleo ainda ficará como matriz energética por um longo tempo, com estimativas de reservas para aproximadamente mais quarenta anos (AGENEAL, 2015). Visto isso, dentre as teorias atualmente presentes para o *demand planning*, ainda há lacunas para modelos que se insiram na produção de petróleo, motivando este modelo preditivo.

### **1.1. Estudo de caso**

Dentre as maiores empresas brasileiras, a PETROBRAS fundada em 1953, ocupa a quarta posição (FORBES, 2016) e no ranking mundial a quadringentésima décima primeira empresa (FORBES, 2016). Mesmo com toda a crise deflagrada pela Operação Lava-Jato e suas implicações na PETROBRAS, ela ainda movimenta e impulsiona boa parte da economia nacional tendo sob sua responsabilidade quase 10% do total de investimentos produtivos do país (EL PAIS, 2016).

Com a desaceleração dos investimentos da PETROBRAS, os fornecedores estão sendo afetados diretamente (DCI, 2014), o que tem levado diversas empresas a revisarem seus investimentos, a se reestruturarem e infelizmente, muitas delas, a encerrarem suas atividades.

O presente trabalho não se aprofundou em estratégias dos governos que controlaram e controlam a PETROBRAS, focando-se apenas no método preditivo para estimativa da produção anual de petróleo. Com esta proposta deste método preditivo, pretende-se agregar e dar mais subsídios no desenvolvimento do plano estratégico dos fornecedores diretos e indiretos que sofrem a influência do mercado petrolífero ou que vislumbram entrar neste mercado.

Para tanto, recorreu-se a dados de mercado e informações diversas em fontes não acadêmicas de forma a dar o respaldo e consistência necessários para o desenvolvimento do presente trabalho, tendo sido realizado o *desk research* de diversos dados apresentados.

### **1.2. Previsão ou Forecasting**

Uma vez definidos a missão e os objetivos de uma determinada empresa, deve-se estabelecer a estratégia para alcançá-los. As análises de mercado e tendências, previsões de vendas, investimentos e custos são a base para elaboração da estratégia das empresas (ALDAY, 2002).

Prever ou o ato de fazer a previsão – o que é corriqueiramente chamado no mundo corporativo de *forecasting* – segundo Lindeke (2005), tem como função primária, prever o

futuro utilizando dados correlacionados com séries temporais ou outros dados disponíveis até os dias atuais, de forma que o resultado destas previsões afete as decisões tomadas hoje.

O *forecasting* é utilizado principalmente para previsões de demandas de produtos e serviços, disponibilidade/ociosidade e necessidade de mão de obra, inventários de materiais e sua necessidade, entre outros. Não deve ser utilizado para excluir informações existentes e, sim, para dar robustez as tomadas de decisões (LINDEKE, 2005).

As previsões de tendências de comportamento de um determinado alvo que se queira prever são sempre um problema dentro da estratégia das empresas. Previsões errôneas infelizmente são o que mais se escuta dentro das organizações: falta de acuracidade com desvios enormes quando se compara o ocorrido em relação as previsões realizadas. E isto comumente visto numa avaliação temporal de curto prazo, o que é bastante crítico, pois quanto maior o prazo de avaliação, maior o risco e comprometimento da acuracidade da previsão.

Previsões mais assertivas precisam considerar a média do comportamento e como a previsão irá oscilar entorno desta média. Isto auxilia a entender o nível de risco que a previsão engloba e o que determinadas oscilações podem acarretar nestas previsões. Isto implica incluir o valor médio e seu desvio *standard*, além do intervalo de acuracidade, com os intervalos com os limites inferior e superior (LINDEKE, 2005).

Portanto, um bom *forecasting* deve ser confiável, ter uma acuracidade tão alta quanto possível, com unidades e parâmetros que façam sentido, além da metodologia ser de fácil utilização e entendimento (LINDEKE, 2005).

### 1.3. Métodos de previsão

Há diversos métodos de previsão, sendo os mais importantes as abordagens qualitativas (ou subjetivas) e as quantitativas (ou objetivas). Considerando os métodos de previsão de demanda e dependendo da complexidade do que se quer prever, pode-se adotar dados de entrada qualitativos e quantitativos.

As abordagens qualitativas ou subjetivas são empregadas quando o processo para analisar os dados não está claramente especificado, podendo ser baseado em experiências do pesquisador, estimativas, opiniões, pesquisas de intenção e julgamentos intuitivos, entre outros (ARMSTRONG, 1985).

Já analisando os métodos ancorados em abordagens quantitativas ou objetivas – que são modelos matemáticos baseados em dados históricos relevantes para o futuro (GAITHER; FRAZIER, 2002, p. 55-86) – há duas abordagens que podem ser seguidas, os modelos causais e os métodos de séries temporais. Os modelos causais estabelecem uma correlação entre variáveis de algo observado ou conhecido com a demanda de algo e as análises de séries temporais observam a demanda histórica de um determinado produto que se quer prever futuramente (BLOCHER ET AL, 2004).

Os tipos de modelo causal mais comumente empregados são a análise de regressão e os modelos econométricos. Para os modelos causais há a correlação na forma  $Y = f(X_1, X_2, X_3, \dots, X_n)$ , onde  $Y$  é a variável que se deseja prever e  $X_i$  são as variáveis que têm o poder de prever  $Y$ . A correlação mais encontrada é a linear na forma  $Y = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n$  (LINDEKE, 2005).

Já os modelos econométricos são geralmente representados por modelos de regressão multivariáveis na forma matricial e sua base reside na aplicação de métodos matemáticos e estatísticos na análise de um conjunto de dados históricos de forma a se prever o dado desejado.

As séries temporais fazem uso de dados do passado para previsões futuras, empregando valores coletados historicamente da variável a ser prevista. O objetivo é isolar padrões nos dados históricos de forma a se obter uma tendência, algum evento que seja cíclico ou sazonal, o que não implica que seja puramente randômico, sem nenhum padrão que se possa observar (LINDEKE, 2005). Segundo Fernandes e Anzanello (2010), estas “séries temporais podem ser representadas por quatro padrões: média, sazonalidade, ciclo e tendência”.

A acuracidade e o risco de cada modelo de previsão elaborado podem ser reduzidos aplicando o modelo obtido em algum dado da base de dados históricos. Desta forma, pode-se calcular possíveis desvios, como o modelo adotado se comporta e qual o risco inerente de possíveis variações na demanda calculada (LINDEKE, 2005). Se a precisão é baixa, o método de cálculo da previsão deve ser alterado.

### 1.3.1. Os modelos quantitativos

Dentre os modelos quantitativos há a regressão linear, onde tenta-se estimar um determinado valor de uma variável Y a partir de valores de outros dados (GAITHER; FRAZIER, 2002, p. 55-86):

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n + \varepsilon$$

sendo,

$\beta_0$ : constante (ponto de intersecção da reta com o eixo y)

$\beta_i$ ,  $i = 1, \dots, n$ : coeficientes da regressão

$X_i$ ,  $i = 1, \dots, n$ : variáveis independentes (dados conhecidos)

$\varepsilon$ : o erro experimental

A regressão linear pode ser simples ou múltipla. Simples quando relaciona somente duas variáveis e múltipla quando relaciona a variável que se quer determinar com um conjunto de n variáveis. Dentro da regressão múltipla, ainda há o método *stepwise forward* e *backward* para cálculo da variável dependente.

Através do método *forward*, onde a partir da variável com maior correlação, será acrescentada uma variável independente de cada vez de forma que se obtenha a maior correlação possível em relação a variável de maior aderência. Desta forma, para as n variáveis independentes escolhidas, pega-se a que teve maior coeficiente de determinação, isto é, maior  $R^2$  ajustado.

Considerando este  $R^2_{ajustado}$  encontrado, adiciona-se uma variável independente de cada vez, verificando-se se o novo  $R^2_{ajustado}$  é maior que o  $R^2_{ajustado}$ . Caso seja maior, adiciona-se esta variável independente testada. Assim sucessivamente até a última variável independente ser testada em relação a variável dependente.

Já a regressão linear múltipla através do método *backward* basicamente faz o caminho inverso do método *forward*, isto é, seleciona todas as variáveis independentes da amostra e vai retirando uma a uma, calculando-se o coeficiente de correlação parcial para a última variável. Então, compara-se com o coeficiente de correlação parcial crítico. Caso seja menor, descarta-

se a variável. Assim é feito até se chegar ao último conjunto de variáveis independentes, sempre descartando a última variável caso o coeficiente de correlação parcial seja menor que o seu valor crítico.

#### **1.4. Variáveis Escolhidas**

Foram escolhidas variáveis macroeconômicas, dados de produção, faturamento e consumo de alguns setores produtivos e alguns dados diversos, conforme descritos a seguir:

##### **1.4.1. PIB – Produto Interno Bruto (valores em reais)**

Como é uma medida do valor de produção, isto é, o PIB pode ser calculado a partir da produção de bens, variações da produção do petróleo refletiriam diretamente em seu cálculo. Nos anos anteriores a 2014, a produção de petróleo no Brasil veio sempre crescendo ano a ano, com investimentos pesados na PETROBRAS, contribuindo substancialmente para o PIB nacional, além de toda a parte de serviços correlatos a área de Óleo & Gás.

##### **1.4.2. Taxa de câmbio (valores em R\$/US\$ para dólar comercial – venda)**

Veríssimo et al. (2012) analisa a relação entre os preços das commodities e a taxa cambial. Um aumento nos preços internacionais das commodities acarretaria num aumento nas receitas de exportação, de forma a ocorrer um aumento da oferta doméstica de moeda estrangeira. Se a demanda for constante, haverá uma valorização da taxa de câmbio real. De forma inversa, se ocorrer uma desvalorização da taxa cambial, haverá aumento dos custos de produção dos bens. Esta seria a relação básica entre a taxa de câmbio e a produção de petróleo.

##### **1.4.3. IED – Investimentos Estrangeiros Diretos (valores em milhões de dólares)**

Analisando o mercado petrolífero, há diversos fatores que poderiam correlacionar a produção de petróleo nacional e o IED, tais como relações de outras multinacionais com a PETROBRAS, integração de operações, por exemplo, da extração e exploração de poços *onshore* e *offshore* até a venda, como a compra de gasolina por outras bandeiras diferentes da PETROBRAS, sua preparação com produtos químicos específicos e venda e necessidade por expansão em outros mercados.

##### **1.4.4. IPCA – Índice Nacional de Preços ao Consumidor Amplo**

Uma relação do IPCA com a produção de petróleo seria sua influência nos preços dos bens e serviços correlatos a sua produção. Mais recentemente, pode-se exemplificar como a política de controle de preços dos derivados de petróleo, maquiando o preço dos combustíveis com preços que não correspondiam ao real custo (impedindo investimentos na PETROBRAS) de forma a aliviar a pressão inflacionária com os congelamentos de tarifas de transportes e pedágios ocorridos anteriormente.

##### **1.4.5. Faturamento líquido da indústria automobilística (valores em milhões de dólares)**

Este setor apresenta forte dependência em relação ao setor petrolífero devido principalmente aos motores movidos à gasolina e diesel. A expansão da indústria automobilística ao longo das décadas gerou a demanda para os combustíveis, mesmo que num ambiente de petróleo

com preços elevados, por exemplo, tenha levado o governo brasileiro a criar o etanol como combustível alternativo a gasolina.

#### **1.4.6. A produção das indústrias naval (valores em Toneladas por Porte Bruto TPB)**

A indústria naval brasileira enfrentou uma grave crise nos anos 90 e, posteriormente, veio se recuperando basicamente devido ao setor petrolífero com a construção de diversas embarcações, de navios sonda e navios de apoio marítimo a plataformas, petroleiros e navios plataforma.

Isto movimentou indústrias de peças e equipamentos, empresas de engenharia, estaleiros, aumentando de forma substancial a criação de empregos no país. Os dados obtidos referentes ao volume de produção nacional contemplam a somatória das encomendas em cada estaleiro do país, obtidos pelo respectivo sindicato do setor (SINAVAL).

#### **1.4.7. A produção de pneus no Brasil (valores em milhares de unidades)**

Como é um item essencial a indústria automobilística e para o setor de transportes, influenciando no custo de bens e serviços e sendo “o segundo custo do transporte rodoviário”, conforme Lagarinhos e Tenório (2008) mencionam em seu trabalho, o consumo de combustíveis, a quantidade de veículos produzidos e o consumo de pneus estão correlacionados.

#### **1.4.8. A produção de asfalto no Brasil (valores em toneladas)**

A produção de asfalto destinada a pavimentação de vias para circulação de automóveis conta com dezessete empresas associadas a ABEDA e com a distribuição do produto para concessionárias, empreiteiras, prefeituras e DER, além de fábricas de emulsões asfálticas (cimento asfáltico disperso em água) e asfaltos modificados também atendendo estes mesmos consumidores. Todo o insumo é proveniente de refinarias, sendo esta a grande relação com o setor petrolífero, alvo do estudo deste trabalho.

#### **1.4.9. Consumo energético de eletricidade (valores em $10^3$ tep)**

Sendo uma importante fonte de energia gerada através de recursos naturais renováveis e não renováveis, tem ligação direta ao objetivo deste trabalho uma vez que:

- para a energia elétrica produzida através de termoelétricas, há a utilização de combustíveis fósseis para a geração de energia, principalmente nos momentos de crises hídricas, como a vivida anteriormente no país;
- além do que indiretamente, para todo o desenvolvimento do país, devido ao grande potencial hídrico brasileiro, há a predominância da utilização de hidrelétricas, alimentando todos os setores de bens e serviços correlatos a indústria do petróleo.

#### **1.4.10. A produção mundial de petróleo (valores em 1000 m<sup>3</sup>/ano)**

A produção mundial de petróleo leva em conta a somatória da produção anual dos continentes africano, americano, europeu e asiático, em especial a região do oriente médio. Foi obtida a série histórica da produção mundial de petróleo para verificação de sua correlação com a produção nacional em termos de tendências mundiais de consumo de petróleo e variações na

produção com aumento ou redução da capacidade ociosa, principalmente por parte dos membros da OPEP.

#### **1.4.11. INCC – Índice Nacional de Custo da Construção**

A princípio, a ideia de se adotar este índice como variável independente foi referente ao mercado imobiliário nas cidades que surgiram ou se desenvolveram graças ao petróleo, como por exemplo, Macaé e Rio de Janeiro, ambos no Estado do Rio de Janeiro e Santos, em São Paulo, que teve um *boom* imobiliário em anos anteriores devido ao pré-sal.

#### **1.4.12. Desembolsos do BNDES para financiamento de investimentos (valores em milhões de reais)**

As liberações realizadas pelo BNDES para financiamento de investimentos no país são destinadas para setores de infraestrutura, agricultura, micro, pequenas e médias empresas, financiamentos à inovação e o setor da indústria, sendo que neste se enquadra os financiamentos as petroquímicas (há outras áreas de financiamento além das citadas).

Como estes investimentos por parte do BNDES visam estimular o crescimento industrial do país, ampliando sua capacidade produtiva e do setor de serviços, aumentando também as exportações e fomentando a capacidade de inovação, por exemplo, através de pesquisas, vem de encontro com as necessidades do setor petroquímico alavancando o aumento de sua produção.

#### **1.4.13. Salário mínimo (valores em reais)**

A série histórica referente ao salário mínimo será considerada a fim de verificar se as oscilações em seus valores se correlacionam a evolução da produção de petróleo no Brasil, como por exemplo, maiores salários, maiores consumos (automóveis, combustíveis, entre outros).

#### **1.4.14. População brasileira**

Com o aumento da população brasileira, haverá o aumento da PEA e se o mercado de Óleo & Gás estiver numa situação de franca ascensão ou ao menos estável, é esperado que haja uma correlação entre a evolução da população brasileira e a produção de petróleo, devido ao consumo de combustíveis e derivados de petróleo que são utilizados como matéria prima para produção de diversos bens.

#### **1.4.15. PEA – População Economicamente Ativa (valores em milhares de pessoas)**

De forma geral, com o desenvolvimento dos setores de bens e serviços, há o aumento da demanda para força de trabalho. Com a evolução da produção de petróleo, ambos os setores foram alavancados refletindo na PEA.

## **2. METODOLOGIA ADOTADA**

Basicamente, o presente trabalho foi dividido em duas frentes, sendo uma a pesquisa dos principais métodos de análise de dados com possíveis modificações e até simplificações do



modelo de forma a se adaptar melhor a análise proposta e a outra, a obtenção destes dados para análise e sua modelagem.

Dentre todos os métodos preditivos pesquisados, por conta das diferentes variáveis escolhidas e o volume de dados coletados para cada variável, adotou-se a Regressão Linear Múltipla como ferramenta matemática para avaliação quantitativa, verificando a correlação e o poder de explicação de cada variável independente em relação a variável dependente – a produção de petróleo em determinado ano.

Já para a coleta de dados, foi elaborada uma pesquisa quantitativa para obtenção dos dados a serem modelados a partir da literatura e *desk research* de forma a se conseguir o maior e mais robusto histórico possível de informações que estivessem relacionadas direta ou indiretamente a demanda de petróleo, a fim de se obter uma regressão dos dados o mais consistente e confiável possível para o método preditivo.

## 2.1. Coleta de dados – inputs para a Regressão Linear Múltipla

Foram coletados dados considerando os anos de 1980 a 2014, com um conjunto com 35 dados para cada variável, volume razoável de dados para uma boa análise. Para alguns dos inputs escolhidos, houve anos em que não foi possível encontrar o respectivo valor, tendo sido realizada a interpolação dos dados secundários obtidos (utilizando os próprios recursos da planilha do EXCEL® da Microsoft) a fim de se obter a série completa no período escolhido.

## 2.2. Regressão Linear Múltipla

A partir da regressão linear múltipla destes dados com o objetivo de se obter como *output* a produção de petróleo no Brasil em um determinado ano, foram testadas as seguintes condições para os inputs escolhidos:

- regressão linear múltipla onde todos os dados foram verificados individualmente de forma a se obter a correlação de cada um com o output, produção de petróleo em determinado ano, obtendo-se a variável independente com maior correlação/aderência;
- regressão linear múltipla através do método *forward*;
- regressão linear múltipla através do método *backward*.

Com os dados coletados e com as condições de testes descritas, o *output* resulta em uma expressão matemática correlacionando as variáveis independentes com a produção de petróleo em determinado ano na seguinte forma:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{11} X_{11} + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{14} X_{14} + \beta_{15} X_{15} + \varepsilon$$

Produção de petróleo em determinado ano =  $\beta_0 + \beta_1$  (PIB) +  $\beta_2$  (taxa de câmbio) +  $\beta_3$  (IED) +  $\beta_4$  (IPCA) +  $\beta_5$  (faturamento líquido da indústria automobilística) +  $\beta_6$  (produção da indústria naval) +  $\beta_7$  (produção da indústria de asfalto) +  $\beta_8$  (produção de pneus) +  $\beta_9$  (consumo energético de eletricidade) +  $\beta_{10}$  (produção mundial de petróleo) +  $\beta_{11}$  (INCC) +  $\beta_{12}$  (desembolsos do BNDES para financiamento de investimentos) +  $\beta_{13}$  (salário mínimo) +  $\beta_{14}$  (população brasileira) +  $\beta_{15}$  (PEA) +  $\varepsilon$

sendo,

- $\beta_0$ : constante (ponto de intersecção da reta com o eixo y)
- $\beta_i$  para  $i = 1, 2, 3, \dots, 15$ : os coeficientes da regressão
- $\varepsilon$ : o erro experimental
- para a regressão linear múltipla onde todos os dados serão verificados individualmente, a expressão esperada será da forma:  
Produção de petróleo em determinado ano =  $\beta_0 + \beta_i$  (variável independente)
- para as regressões lineares múltiplas através dos métodos *backward* e *forward*, a expressão será composta por duas a quinze variáveis independentes.

Como para cada modelagem obtida é calculado o coeficiente de determinação  $R^2$ , pode-se então verificar a aderência do modelo em relação a variável dependente. O  $R^2$  é um indicativo de quanto o modelo consegue explicar os valores observados, isto é, quanto maior o  $R^2$ , mais aderente é o modelo calculado, de forma que a variável dependente consegue se ajustar melhor à amostra, de uma ou mais variáveis independentes.

Portanto, o  $R^2$  de um determinado modelo significa o quanto a variável dependente consegue ser explicada pelas variáveis independentes escolhidas no modelo (o  $R^2_{ajustado}$  leva em consideração as diferentes unidades de cada variável independente utilizada na amostra).

### **3. APLICABILIDADE E ANÁLISE DOS RESULTADOS**

#### **3.1. Resultado das simulações**

A partir dos inputs contemplando a respectiva série histórica entre os anos de 1980 a 2014, foram realizadas as regressões lineares múltiplas simulando todas as variáveis independentes separadamente e juntas, no método *backward* e no *forward*, utilizando o pacote estatístico SPSS versão 8.0 (lembrando que quanto maior o coeficiente de determinação  $R^2$ , mais explicativo e mais aderente é o input em relação ao output e para valores de  $R^2$  acima de 70% os inputs são considerados muito bons para o modelo escolhido).

##### **3.1.1. Análise de cada variável separadamente**

Com as quinze variáveis independentes selecionadas, foi realizada a regressão linear múltipla do conjunto obtendo os resultados da Tabela 1. Observa-se que das quinze variáveis independentes selecionadas, sete tiveram alta aderência com o output, com coeficiente de determinação acima de 76%:

VARIÁVEIS INDEPENDENTES	R <sup>2</sup>
Produção de Pneus	90,60%
PEA	90,40%
População Brasileira	89,90%
Produção Mundial de Petróleo	88,00%
Consumo Energético - Eletricidade	87,80%
Faturamento Líquido do Setor Automobilístico	78,10%
Produção de Asfalto	76,50%
Taxa de Câmbio	68,60%
Desembolsos do BNDES	63,30%
IED	49,20%
IPCA	33,00%
Produção da Indústria Naval	22,40%
Salário Mínimo	15,60%
INCC	—
PIB	—

**Tabela 1: Resultados individuais da simulação**

Esperava-se certa correlação do INCC com o *output* devido a logística envolvida na construção civil com o consumo de combustíveis e com o mercado imobiliário relacionado com o desenvolvimento da infraestrutura habitacional devido ao mercado de Óleo & Gás. Também lembrando outros fatores que foram considerados como *inputs* como o mercado automobilístico com a aquisição de automóveis e peças, a contribuição do setor civil no PIB, a geração de renda e o aumento do número de trabalhadores na construção civil favorecendo a PEA, acreditava-se haver alguma correlação, o que, no entanto, não houve relevância para este *input*.

O mesmo ocorreu para o PIB, onde também não houve relevância em relação ao *output*, o que de certa forma, considerando que o PIB pode ser calculado a partir da produção de bens e serviços e a produção de petróleo veio aumentando constantemente ao longo dos anos, esperava-se que variações da produção do petróleo deveriam ter refletido ao menos um pouco no *output*.

Por conseguinte, das hipóteses de aderência de ambos os *inputs* em relação ao *output* não há relação alguma das hipóteses levantadas e não servem como variável independente para determinação do *output*.

Portanto, para a produção de petróleo no Brasil em um determinado ano:

- o a melhor variável independente a ser escolhida é a produção de pneus, por ter um poder de explicação de 90,60%, considerando que referente a sua amostra de 35 dados, somente 6 deles foram interpolados.

$$\text{Produção de petróleo em determinado ano} = - 29830,018121124 +$$

$$+ 2,24661032734761 (\text{produção de pneus})$$

- o logo a seguir, vem a população brasileira e a PEA, que são dados fáceis de se encontrar e com estimativas constantemente elaboradas por parte do governo. Como também apresentam alto poder de explicação para o *output*, distando menos de 1% do R<sup>2</sup> da produção de pneus e considerando que somente a PEA teve 7 amostras interpoladas em 35, são duas variáveis independentes muito recomendadas para utilização.

$$\text{Produção de petróleo em determinado ano} = -75293,1568839717 + 1,88938992526928 (\text{PEA})$$

- o apesar da produção de asfalto ter tido alto poder de explicação, como houve muitos dados interpolados utilizados nas regressões (18 em 35 dados) e pela facilidade de se encontrar dados relativos aos outros inputs com  $R^2$  maior que da produção de asfalto, recomenda-se não considerá-lo.
- o caso encontre-se dados relativos a possíveis previsões da produção mundial de petróleo e do consumo energético brasileiro, pode-se também utilizar como inputs, lembrando que são dados mais difíceis de se encontrar do que PEA e crescimento populacional brasileiro.

OBSERVAÇÃO: não serão apresentadas todas as funções que relacionam o output às variáveis independentes, somente as principais com maior  $R^2$ .

### 3.1.2. Análise de possíveis composições destas variáveis

Já para os métodos de regressão *forward* e *backward*, a fim de se obter o maior coeficiente de determinação para o respectivo modelo estatístico, foram obtidos os seguintes resultados:

#### 3.1.2.1. Método de Regressão *Forward*

No caso estudado, o modelo mais “otimizado” contemplou as variáveis independentes “produção de pneus” e “produção da indústria naval”:

Modelo	R	$R^2$	$R^2$ AJUSTADO	DESVIO PADRÃO DA ESTIMATIVA
1	0,952027940146859 <sup>a</sup>	0,906357199	0,903519538	11260,67988
2	0,957860364098493 <sup>b</sup>	0,917496477	0,912340007	10733,60581

a VARIÁVEIS INDEPENDENTES: (Constante), Produção de Pneus

b VARIÁVEIS INDEPENDENTES: (Constante), Produção de Pneus, Produção da Indústria Naval

Modelo		coeficientes não uniformizados	desvio padrão	coeficientes uniformizados <sup>a</sup>	t	Sig.
		B		Beta		
1	(Constante)	-29830,01812	5540,9915		-5,383516311	6,0829E-06
	PNEUS	2,246610327	0,125706659	0,95202794	17,87184816	2,1335E-16
2	(Constante)	-34062,08029	5660,487107		-6,017517512	9,85018E-07
	PRODUÇÃO DE PNEUS	2,427123354	0,147984542	1,02852249	16,4011952	2,48191E-16
	PRODUÇÃO DA INDÚSTRIA NAVAL	-0,006259162	0,003011265	-0,130348358	-2,078582514	0,045756467

VARIÁVEL DEPENDENTE: PETRÓLEO

**Tabela 2: coeficiente de determinação e coeficientes do modelo linear para o método de regressão *forward***

O coeficiente de determinação foi de 91,7% e o ajustado de 91,23%, sendo o modelo estatístico encontrado:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

$$\text{Produção de petróleo em determinado ano} = -34062,080291857 + 2,4271233544225 (\text{produção de pneus}) - 0,00625916203343831 (\text{produção da indústria naval})$$

Individualmente conforme apresentado na Tabela 2, a “produção de pneus” tem um poder de explicação de 90,6% e a “produção da indústria naval” de apenas 22,4%. No entanto, quando ambas as variáveis independentes são analisadas em conjunto, há um ganho de quase 0,7% para o  $R^2$ .

Apesar de ser um incremento baixo e a variável “produção da indústria naval” pouco explicar a relação com o output, pode ter ocorrido um efeito de sinergia entre ambas as variáveis resultando no melhor par de variáveis independentes para explicar o modelo preditivo.

Como além da variável “produção de pneus” havia mais outras seis variáveis independentes com alto poder de explicação (acima de 70% de poder de explicação), não ocorreram mais correlações com outras variáveis muito provavelmente por já serem muito próximas, isto é, correlação muito alta com comportamento muito parecido.

Por conseguinte, do conjunto de doze variáveis, sobraram quatro com  $R^2$  entre 16% e 49%. Dentre estas quatro variáveis independentes, a que possui relação explicitamente direta com a produção de pneus e petróleo é a indústria naval. A maior demanda do setor naval é o mercado de Óleo & Gás e para toda a logística de bens e serviços necessária para a execução dos contratos, há a necessidade de pneus e combustível como insumos para o transporte. Vale ressaltar que mesmo sendo a melhor combinação encontrada e ao se adicionar mais uma variável, comparando-se individualmente só houve o incremento de 0,7%, o que é ínfimo para estimativa numa ordem de 91%, poderia ainda sim, somente levar em consideração a produção de pneus. Outro ponto é a grande quantidade de dados interpolados para a variável “produção da indústria naval”.

### 3.1.2.2. Método de Regressão *Backward*

O melhor modelo, cujo  $R^2$  ajustado foi de 92,8%, apresentou apenas quatro inputs: faturamento líquido do setor automobilístico, produção de pneus, população brasileira e consumo energético – eletricidade. Apesar de não ser o modelo com maior poder de explicação, pois o modelo melhor obteve  $R^2$  de 94,3%, ele apresenta apenas 4 variáveis ao invés dos 15 inputs, o que o torna um modelo muito mais simples com  $R^2$  de 93,6%, sendo a diferença em termos de poder de explicação de apenas 1%, o que é um ganho baixo.

Como o  $R^2$  ajustado já leva em consideração o número de variáveis, quanto menos variáveis a se trabalhar melhor. Então, em quatro modelos foram removidas as variáveis macroeconômicas IPCA, PIB, taxa de câmbio e IED, indicando pouca relevância destes fatores na produção de petróleo. Isto já era sinalizado nos resultados das simulações anteriores, onde na análise individual de cada input, o PIB não teve relevância alguma para a modelagem, IPCA e IED tiveram baixo poder explicativo e somente a taxa de câmbio se aproximou do limite de 70% em relação ao coeficiente de determinação. O mesmo ocorreu no método *forward*, onde novamente os índices macroeconômicos não tiveram relevância.

Dois fatores bastante ligados a produção nacional de petróleo, que são as variáveis “produção da indústria naval” e “produção mundial de petróleo”, também foram excluídas do modelo final e também não constavam na regressão *forward*, inclusive na análise individual, a produção da indústria naval apresentou pouca relevância. Mesmo que a produção nacional de petróleo siga a tendência do mercado mundial de petróleo e a indústria naval tenha forte dependência dela em relação ao setor de Óleo & Gás, não foi suficiente para estarem no modelo mais otimizado.

Já para o INCC, mais uma vez não teve relevância em nenhuma das modelagens, lembrando que seu poder de explicação quando analisado individualmente já era irrelevante. Isto

corroborar a ausência de aderência deste índice em relação a alguma possibilidade de relação com o output.

Para os desembolsos do BNDES para diversos investimentos, também não teve relevância no modelo *backward*, vendo que seu poder explicativo quando analisado individualmente foi mediano e também não constou do modelo *forward*.

Observando a produção nacional de asfalto e a PEA, que foram inputs com alta relevância quando analisados individualmente, também não foram considerados no modelo mais otimizado, sendo que o salário mínimo que teve a correlação individual mais baixa foi a variável que mais se aproximou do modelo adotado para a regressão *backward*.

Portanto, pelos resultados obtidos, o modelo preditivo a ser adotado é o que se segue:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$$

Produção de petróleo em determinado ano = -200145,0685 + 0,67681341 (faturamento líquido da indústria automobilística) + 1,826830151 (produção de pneus) - 4,530430405 (consumo energético de eletricidade) + 0,001649641 (população brasileira)

### 3.1.2.3. Análise Geral

O aumento do poder preditivo dos modelos encontrados, o que significa maior poder de explicação para as estimativas da produção nacional de petróleo para um determinado ano, implicou em se utilizar modelos com mais variáveis.

Caso se entenda que 91,7% de aderência ao output já é um bom poder de explicação, pode-se adotar o método *forward* com o levantamento da produção de pneus e da indústria naval, lembrando que um ponto negativo para a produção da indústria naval é a dificuldade em se encontrar dados históricos, o que acarretou na necessidade de se interpolar grande quantidade de dados.

Mas, se o aumento do poder de explicação for um fator determinante, pode-se adotar o método *backward* tendo como primeira opção com  $R^2$  de 93,6% o modelo com quatro inputs e como segunda opção com 94,3% o modelo com 15 inputs, salientando que há um baixo incremento do poder de explicação de apenas 1% entre ambos os modelos aliada a dificuldade de se coletar todos os 15 inputs e com a possibilidade de se realizar grande número de interpolações para os dados inexistentes.

Contudo, para uma estimativa rápida e com uma aderência relativamente alta com mais de 89% podem ser utilizados os dados de população brasileira ou PEA, pois são dados fáceis de se obter e que o governo sempre atualiza e projeta para anos futuros, lembrando que são os resultados da regressão linear individual para cada input. Um pouco mais dificultoso, porém com maior poder de explicação vem a produção de pneus, que eleva o  $R^2$  para 90,6%.

Portanto, estimar 11 ou mais variáveis para ganhar somente 4% de aderência no modelo é um ganho muito baixo, sendo que há de 89% a 94% de chance de acerto na previsão da produção de petróleo com uma ou até quatro variáveis independentes. Mesmo que haja bilhões de dólares em investimentos conforme o plano de investimentos da PETROBRAS, o máximo que se poderia alcançar seria 94% de poder de explicação.

#### 4. CONSIDERAÇÕES FINAIS

Apesar dos métodos de previsão atualmente existentes, previsões de longo prazo são ainda muito complexas e podem incorrer em diversos erros. A sensibilidade dos responsáveis em fazer os *forecasts* em relação ao setor em que se utilizará a previsão da demanda também é de suma importância para poder interpretar o resultado e validá-lo de forma a atenuar possíveis fatores externos que possam interferir no *forecast*.

O envolvimento de outras áreas, como, execução (engenharia, gerenciamento de projetos e suprimentos e logística), comercial e vendas e financeira e econômica para fazerem suas considerações e tendências do mercado também são de bastante valia para a análise dos resultados.

Um fator muito importante e que o modelo apresentado não consegue levar em consideração é o fator político e econômico. Crises no governo como pode ser mencionada a atual crise da PETROBRAS com a Operação Lava Jato e no próprio setor petrolífero, como a crise no preço do barril do petróleo aparentemente deflagrada pelos países do Oriente Médio em relação a concorrências com o gás de xisto e problemas com a Rússia, o que acarretou na derrubada do preço do barril cotado a mais de cem dólares até meados de agosto de 2014 para menos de sessenta dólares em março de 2015 (INVESTING.COM, 2015).

Como o porte do mercado brasileiro para o setor petrolífero é bastante expressivo e as perspectivas futuras sinalizam mudanças na política e na forma de condução dos negócios da PETROBRAS, há a possibilidade de se ter um ambiente de mercado mais ágil, competitivo e transparente, onde o planejamento da demanda para todas as áreas que circundam o mercado de óleo e gás voltará a ter relevância.

#### 5. REFERÊNCIAS

AGENEAL. *Homepage* institucional. Disponível em: <http://www.ageneal.pt/content01.asp?BTreeID=00/01&treeID=00/01&newsID=7>. Acesso em diversas datas entre dezembro de 2014 e março de 2015.

ALDAY, Hernan E.C.. **Estratégias Empresariais Coleção Gestão Empresarial**. FAE Business School, 2002.

ARMSTRONG, J. Scott, **Long-Range Forecasting: From Crystal Ball to Computer**, Wiley-Interscience; 2 edition (July 4, 1985).

BLOCHER, James D., Mabert, Vincent A., Soni, Ashok K., Venkataramanan, Munirpallam A. **Forecasting, Including an Introduction to Forecasting using the SAP R/3 System**. Indiana University Kelley School of Business, February 2004.

DCI. Disponível em: <http://www.dci.com.br/industria/demanda-fraca-afeta-fornecedor-da-petrobras-id406853.html>. Acesso em diversas datas entre dezembro de 2014 e março de 2015.

EDITORA ABRIL, PLANETA SUSTENTÁVEL. Disponível em: [http://planetasustentavel.abril.com.br/noticia/energia/conteudo\\_280677.shtml](http://planetasustentavel.abril.com.br/noticia/energia/conteudo_280677.shtml). Acesso em diversas datas entre dezembro de 2014 e março de 2015.

EL PAIS ECONOMIA. Disponível em: [http://brasil.elpais.com/brasil/2016/01/12/economia/1452625037\\_799572.html](http://brasil.elpais.com/brasil/2016/01/12/economia/1452625037_799572.html). Acesso em diversas datas entre junho e julho de 2016.

FERNANDES, Fabrício e ANZANELLO, Michel J., **Integração de métodos quantitativos e qualitativos para previsão de demanda no setor de autopeças**, UFRGS, 2010.

FORBES. Disponível em: <http://www.forbes.com.br/listas/2016/05/19-maiores-empresas-do-brasil-em-2016/#foto16>. Acesso em diversas datas entre junho e julho de 2016.

GAITHER, Norman e FRAZIER, Greg. **Administração da Produção e Operações**. 8 edição. São Paulo, Thomson Pioneira, p. 55-86, 2002.

INVESTING.COM. Disponível em: <http://br.investing.com/commodities/brent-oil-historical-data>. Acesso em diversas datas entre dezembro de 2014 e março de 2015.

LAGARINHOS, Carlos Alberto F. e TENORIO, Jorge Alberto S.. **Tecnologias utilizadas para a reutilização, reciclagem e valorização energética de pneus no Brasil**. Polímeros (online), 2008.

LINDEKE, Richard R., **Forecasting Models – Chapter 2**, Professor of Mechanical & Industrial Engineering, Department of Mechanical & Industrial Engineering, University of Minnesota-Duluth, Mn, 2005. Disponível em: [http://www.slidefinder.net/f/forecasting\\_models\\_chapter\\_3265\\_lindeke/ie3265\\_forecasting/28373036](http://www.slidefinder.net/f/forecasting_models_chapter_3265_lindeke/ie3265_forecasting/28373036). Acesso em diversas datas entre dezembro de 2014 e março de 2015.

PROMINP. *Homepage* institucional. Disponível em: [http://www.prominp.com.br/prominp/pt\\_br/conteudo/conteudo-local.htm](http://www.prominp.com.br/prominp/pt_br/conteudo/conteudo-local.htm). Acesso em diversas datas entre dezembro de 2014 e março de 2015.

VERÍSSIMO, M. P.; XAVIER, C. L. & VIEIRA, F. (2012). **Taxa de Câmbio e Preços de Commodities: Uma Investigação sobre a Hipótese da Doença Holandesa no Brasil**. Revista Economia.