



**UNIVERSIDADE ESTADUAL DE CAMPINAS**  
**FACULDADE DE CIÊNCIAS APLICADAS**



VAGNER YUJI HORITA

## **Determinantes dos vieses de previsão do preço futuro do algodão**

Orientador:

Prof. Dr. Antonio Carlos Moretti

Co-orientador:

Prof. Dr. Johan Hendrik Poker Jr

Limeira - SP  
2014



**UNIVERSIDADE ESTADUAL DE CAMPINAS**  
**FACULDADE DE CIÊNCIAS APLICADAS**



## **Determinantes dos vieses de previsão do preço futuro do algodão**

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do título de bacharel em Engenharia de Produção à Faculdade de Ciências Aplicadas da Universidade Estadual de Campinas.

Orientador:  
Prof. Dr. Antonio Carlos Moretti  
Co-orientador:  
Prof. Dr. Johan Hendrik Poker Jr

Limeira - SP  
2014

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA PROF. DR. DANIEL JOSEPH HOGAN  
DA FACULDADE DE CIÊNCIAS APLICADAS

H782d Horita, Vagner Yuji  
Determinantes dos vieses de previsão do preço futuro do algodão / Vagner Yuji  
Horita. - Limeira, SP: [s.n.], 2014.  
29 f.

Orientador: Antônio Carlos Moretti.  
Co-orientador: Johan Hendrik Poker Junior.  
Monografia (Graduação) – Universidade Estadual de Campinas, Faculdade de  
Ciências Aplicadas.

1. Preço. 2. Algodão. 3. Previsão. 4. Métodos de previsão. 5. ANOVA. I.  
Moretti, Antônio Carlos. II. Poker Junior, Johan Hendrik. III. Universidade  
Estadual de Campinas. Faculdade de Ciências Aplicadas. III. Título.

Título em inglês: Determinants of bias forecasting the future price of cotton

Keywords: - Price;

- Cotton;

- Forecast;

- Forecasting methods;

- ANOVA.

Titulação: Bacharel em Engenharia de Produção.

Banca Examinadora: Prof. Dr. Antônio Carlos Moretti.

Prof. Dr. Johan Hendrik Poker Junior.

Prof. Dr. Aníbal Tavares de Azevedo.

Data da defesa: 11/07/2014.

## Resumo

Trabalho realizado com intuito de mostrar a influência das variáveis divulgadas pelo USDA (Departamento de Agricultura dos Estados Unidos), na variação das cotações do algodão em pluma e verificar a acurácia dos métodos de previsão nos quartis extremos, em relação ao valor real verificado no dia do fechamento do contrato.

Inicialmente esperava-se que os quartis extremos, primeiro e o quarto, possuísem diferenças mais significativas em relação ao quartis intermediário, dado que os modelos preditivos utilizam como base uma distribuição gaussiana normal, que possui uma tendência de subvalorizar os extremos inferiores e supervalorizar os superiores.

Entretanto, após rodar análises do teste T de Student para diferença de média no software Gretl, verificou-se que o erro entre os quartis não era significativo a ponto de se afirmar que os modelos possuem erros mais acentuados nos quartis extremos.

No caso das variáveis, foi realizada uma ANOVA (Análise de variância) para verificar qual dos parâmetros possui maior conexão com a movimentação dos preços da pluma de algodão na bolsa de Nova Iorque. Foi identificado que a variável "produção" em seu quarto quartil gera vieses significativos na previsão do preço do algodão.

## **Abstract**

This paper was performed in order to show the influence of variables reported by the USDA (U.S. Department of Agriculture), in the variation of the prices of cotton lint and verifies the accuracy of forecasting methods in extreme quartiles, relative to the real price on the day that the contract ends.

Initially it was expected that the ends first and the fourth quartiles possessed most significant differences from the intermediate quartiles, since predictive models using as a basis a standard Gaussian distribution, which has a tendency to underestimate the extreme lower and upper overvalue.

However, after running the analysis Student t test for mean difference in Gretl software, it was found that the error between quartiles was not significant to the point of stating that the models have more severe errors in the extreme quartiles.

In the case of variables, an ANOVA was performed to determine which parameters have greater connection with the movement of prices of raw cotton on the New York Stock Exchange.

## Sumario

1 - Introdução .....	8
2 - Revisão Bibliográfica .....	10
2.1 - Modelos de Previsões .....	10
2.1.1 - Séries temporais – ARIMA.....	10
2.1.2 - Modelos computacionais .....	11
2.1.2.1 Redes Neurais.....	11
2.2 - Teoria Comportamental – Excesso de confiança .....	14
2.3 - Distribuição para Valores Extremos .....	15
3 - Metodologia .....	16
3.1 – Dados.....	16
3.1.1 – Dados dos Preços .....	17
3.1.2 – Dados das Variáveis.....	18
3.2 – Análise de dados.....	19
4 – Resultados .....	20
5 – Conclusão .....	29
6 - Referencia Bibliográfica .....	30

## 1 - Introdução

Com o crescimento da população mundial e o aumento do consumo de alimentos, as áreas relacionadas à produção de grãos e outras commodities agrícolas ganharam notoriedade e destaque de grandes investidores.

O que antes era conhecido como trabalho rural subdesenvolvido, agora se torna extremamente profissional e tecnológico, mirando altas produtividades para suprir a demanda mundial.

Uma das responsáveis pelo grande aumento do consumo, notado na última década, foi a entrada da China como potência econômica e maior importadora de grãos e algodão do mundo. Com a melhora da qualidade de vida da população chinesa, o consumo por alimentos aumentou de forma estrondosa, tanto em quantidade quanto em qualidade. Os chineses passaram a consumir mais carne em comparação às décadas passadas, gerando um maior consumo de ração animal, que tem como matéria-prima, proteína vegetal.

A profissionalização do setor, fez com que os produtores e compradores não ficassem a mercê do mercado e fixassem seus ganhos e perdas na bolsa de valores, transferindo o risco da variação do preço para os especuladores, que dão liquidez às negociações.

De acordo KOTLER (1991) e MOON *et. al*, (1998) quanto mais longo forem os prazos dos planejamentos, maior será a dificuldade de sua elaboração, devido a incertezas no setor econômico, sócio-político e tecnológico. Por exemplo, as flutuações nos mercados de capital, logística, mão-de-obra e vendas, ajudam na concepção de um ambiente incerto e extremamente competitivo. Atualmente as empresas estão dando grande atenção com o que pode acontecer e como atuar sobre esses eventos preventivamente, ou como adaptar suas estratégias às mudanças previstas, dado que a situação do mercado se encontra em constante evolução e apresenta um dinamismo nunca antes vivenciado.

Entretanto, grandes problemas ainda preocupam os participantes da bolsa como: Qual a tendência do mercado? Qual o valor do produto no dia do fechamento do contrato?

Perguntas que são difíceis de responder, pois os métodos de previsão existentes não estão aptos a definir com exatidão qual será o patamar a ser negociado, e pelos históricos de negociações coletados, quanto maior a variável tempo, maior é o erro encontrado em relação ao preço final da mercadoria no dia do encerramento do contrato em relação à previsão realizada.

Como grande parte dos meios utilizados para estimativa de preço são baseados em uma distribuição gaussiana normal, tende-se a notar uma subvalorização nos valores localizados no primeiro quartil da distribuição, assim como uma supervalorização dos nos valores pertencentes ao quarto quartil, ou seja, os valores situados nos extremos são muito distantes dos valores verificados na realidade. Sendo assim, pode-se verificar uma tendência quando os valores do produto estão abaixo do segundo quartil e acima do terceiro quartil.

Assim, com as informações divulgadas pelas agências de notícias, cabe aos participantes da bolsa, analisá-las de um modo favorável, para tentar prever o comportamento do mercado, se o sentimento é altista (bullish), tendência de alta nos preços ou baixista (bearish), tendência de baixa nos preços. Análise conhecida como fundamentalista, visto que utiliza fundamentos divulgados pelo setor para identificar a direção que caminham os negócios realizados.

## **2 - Revisão Bibliográfica**

### **2.1 - Modelos de Previsões**

Segundo ARMSTRONG (1998) os métodos de previsão têm sido desenvolvidos com o fim de antecipar estados futuros, de fatores e/ou variáveis que interfiram de algum modo no planejamento estratégico para garantir uma produção ou atividade sem grandes surpresas e com custo.

Na atividade agropecuária, a necessidade de informações preditivas de variáveis de mercado é fundamental, dada a defasagem existente entre as decisões de produção e seus efeitos (VERE e GRIFFITH, 1990).

O agronegócio se depara com diversos tipos de riscos, tais como: preço, clima, crédito, custos, operacional, entre outros. E umas das variáveis mais importantes para o sucesso do negócio é o preço da commodity. Desta forma, um gerenciamento eficiente do risco de preço é fundamental para que a atividade de longo prazo seja sustentável e lucrativa, e muitos instrumentos existem no mercado para auxiliar nesta decisão (EGELKRAUT et al., 2006).

Dentre os modelos mais utilizados estão os que levam em consideração séries temporais com modelos univariados que possuem acesso mais fácil e custo reduzido, porém não levam muitas variáveis em consideração, os modelos computacionais que utilizam diferentes softwares e são mais usados por bancos e grandes investidores e possuem o custo mais elevado, porém conseguem analisar diversas variáveis simultaneamente e com maior precisão.

#### **2.1.1 - Séries temporais – ARIMA**

A análise de séries temporais se aplica nos casos em que há um padrão persistente ou sistemático no comportamento da variável, que é possível de captar através de uma representação paramétrica (PINDYCK e RUBENFIELD, 1991).

Muitas séries temporais econômicas mudam suas características estocásticas ao longo do tempo de observação. São as chamadas séries não estacionárias. Estas

séries podem, no entanto, exibir algum tipo de homogeneidade. Ou seja, podem se comportar de forma homogênea aparte de níveis locais e/ou tendência (Box, Jenkins e Reinsel, 1994).

Segundo Pindyck & Rubenfield (1991) os modelos ARIMA (Autorregressivo – Integrado – Média móvel), inicialmente formulados por Box & Jenkins (1976), são baseados no pensamento de que é possível se modelar uma série temporal não estacionária a partir de  $d$  diferenciações e do auxílio de um componente autorregressivo e de um componente média móvel.

### **2.1.2 - Modelos computacionais**

Os modelos computacionais surgiram para auxiliar a ciência e a matemática na análise de dados, visto que os novos problemas possuíam uma quantidade tão elevada de informações e variáveis, que o cérebro humano não era mais capaz de processá-las e verificar a existência de padrões ou repetições em determinados casos.

Mesmo com todos os avanços técnicos, científicos e a introdução da inteligência artificial, os erros das predições continuam inevitáveis, de acordo com EVANS (1982). Pois o procedimento segue limitado pelas suposições atribuídas a ele.

#### **2.1.2.1 Redes Neurais**

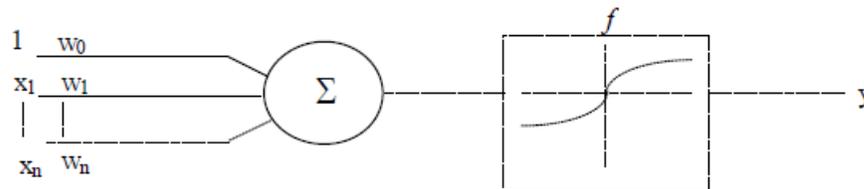
De acordo MACCULLOCH e PITTS (1943), Warren McCulloch, um psicólogo e Walter Pitts, um matemático foram os primeiros a realizar trabalhos na área de Redes Neurais Artificiais (RNA). Eles introduziram os conceitos de combinação de neurônios, pesos e *thresholds* na análise de séries temporais e amostras de dados.

Os modelos de RNA se diferenciam dos modelos tradicionais de previsão por serem modelos não paramétricos, envolvendo algoritmos de aprendizado. Tais algoritmos buscam imitar a estrutura de interconexões do cérebro humano, com o intuito de incorporar o padrão de comportamento de uma série temporal de modo a

prever, da maneira mais eficiente possível, valores futuros dessa série (TURBAN, 1993).

De acordo com ZIRILLI (1996) e TURBAN (1993), as RNA são formadas por vários neurônios artificiais independentes e capazes de interagir entre si, funcionando como pequenos processadores.

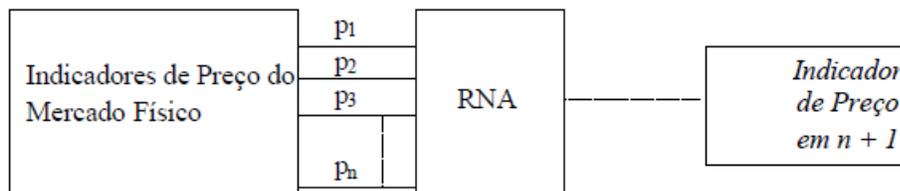
A unidade básica de uma RNA é o neurônio artificial (Figura 1). Sua estrutura busca reproduzir a estrutura do neurônio humano, tendo, assim, três componentes básicos: um elemento somatório  $S$ , uma função de ativação  $f$  e as ponderações  $w_i$ .



Fonte: ZIRILLI (1996)

**Figura 1 – O neurônio artificial**

A grande vantagem desse modelo está na capacidade de se desenvolver através de acontecimentos prévios, exemplos e melhorar seu desempenho se adaptando às mudanças do ambiente. Além disso, pode lidar com informações incompletas e dados defeituosos, obtendo resultados satisfatórios através da generalização de padrões de treinamento.



Fonte: ZIRILLI (1996)

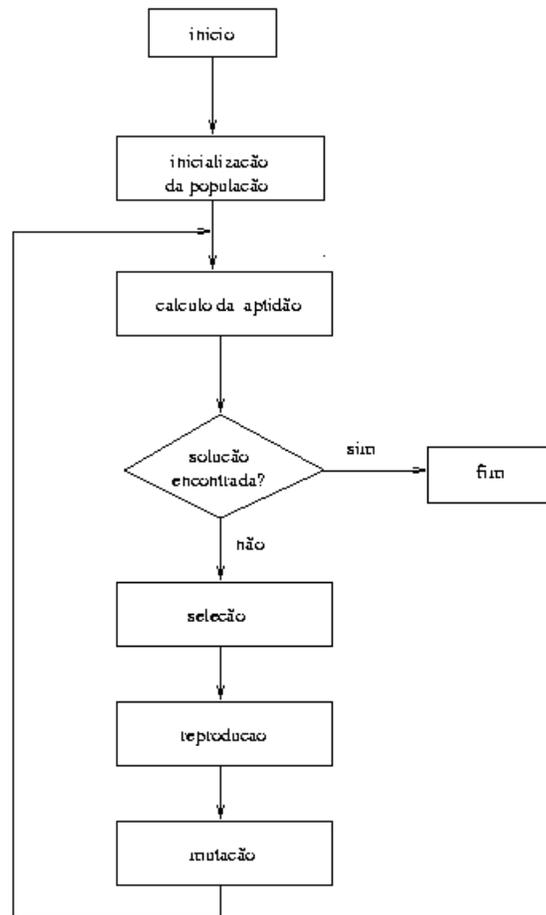
**Figura 2 – Esquema de uma RNA para previsão de preços**

### **2.1.2.2 Algoritmo Genético**

Os algoritmos genéticos são modelos de máquinas de aprendizagem que usam a metáfora da evolução das espécies numa perspectiva de simulação desses processos com o fim último de ajudar a resolução de problemas têm como objetivo aperfeiçoar um conjunto de parâmetros que instanciados nos dão a solução dos problemas.

Em princípio os algoritmos genéticos podem modelar qualquer função que seja computacionalmente representável. Entretanto esses tipos de modelos são ultrapassados em termos de desempenho em situações em que os procedimentos de resolução de problemas se encontram completamente sistematizados.

O ponto forte desse método é o de se ajustar conforme as necessidades do problema e evoluir através de indicação de novos parâmetros ou interpolações realizadas pelo próprio sistema. No caso das previsões de preços, o algoritmo se ajusta de acordo com as diferenças de médias entre o preço previsto e o preço real no dia do vencimento, podendo realizar modificações e evoluir de acordo com o resultado obtido em uma das suas etapas.



Fonte: UFRJ

**Figura 3 – Estrutura básica de um Algoritmo Genético**

## **2.2 - Teoria Comportamental – Excesso de confiança**

DE BONDT e THALER (1995) sugerem que provavelmente uma das descobertas mais importantes da psicologia comportamental seja a de que as pessoas são extremamente confiantes, podendo muitas vezes afetar seu julgamento e influenciar diretamente em algumas decisões.

Segundo COPELAND e WESTON (1988), após a grande mudança proporcionada por Markowitz (1952), na área de mercado de capitais, os pensadores do ramo voltaram suas atenções para o funcionamento do mercado e para o comportamento dos investidores. Na teoria, espera-se sempre que suas

movimentações sejam perfeitamente eficientes, nas quais os preços reflitam todos os fatores informacionais e não existam informações relevantes em poder de uns ou outros indivíduos. Sendo assim, a única pressuposição é a de que os investidores busquem apenas a maximização do possível retorno do seu portfólio.

De acordo com NOFSINGER (2006), o excesso de confiança influencia diretamente no processo decisório fazendo o investidor acreditar que possui informações privilegiadas em relação ao mercado, melhores condições de selecionar investimentos com maior rentabilidade e o faz crer também que seus ativos estão livres dos riscos que rondam os demais. Ou seja, espera-se que o indivíduo possua excesso de confiança superestime suas capacidades e subestime os riscos, um comportamento arriscado ao se tratar de mercado de capitais.

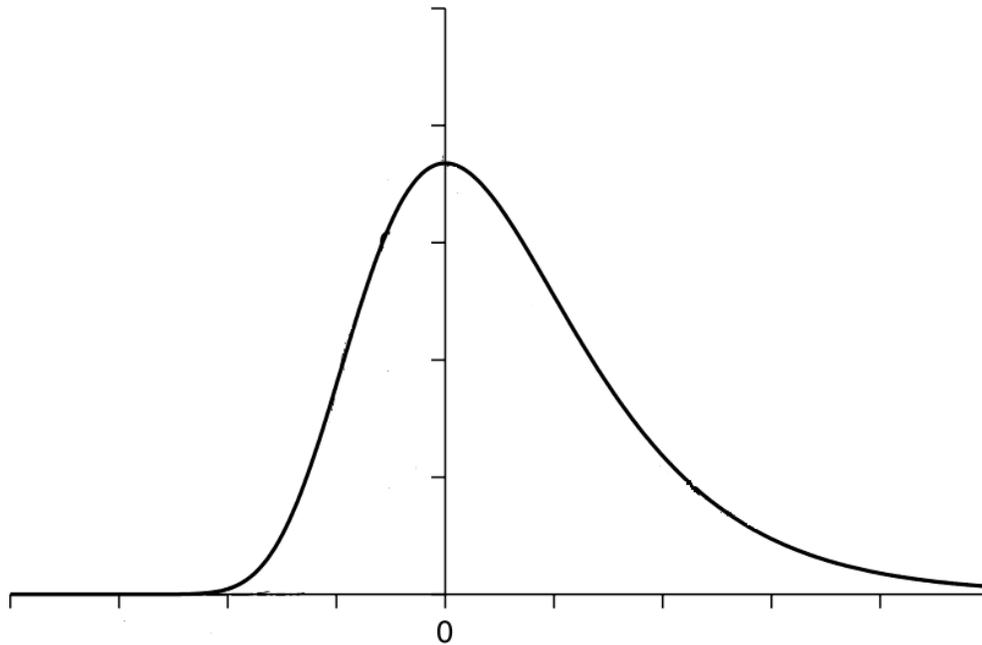
No mercado de ações e mercadorias esse comportamento de excesso de confiança, faz com que produtores sempre supervalorizem seus produtos e peçam valores acima do praticado nos pregões, causando uma distorção na curva do preço e algumas vezes forçando alguns compradores a adquiri-la mesmo acima da real cotação. Esse fator emocional, ligado aos participantes do mercado e tomadores de decisões, dificulta ainda mais a leitura das cotações e torna a previsão de preços uma tarefa ainda mais complicada e quase impossível.

### **2.3 - Distribuição para Valores Extremos**

FISHER e TIPPETT (1928), primeiros a formalizar a teoria para distribuição de valores extremos, definiram três tipos possíveis de distribuições (I, II e III), que posteriormente também seriam conhecidas como Gumbel, Fréchet e Weibull.

Teorema que está relacionado ao estudo de acontecimentos raros, como catástrofes naturais, cálculo de seguros e eventos pouco comuns no mercado financeiro, funciona tanto para cálculo de valores máximos quanto para valores mínimos. O mais usado de mercado financeiro é o modelo de Gumbel, que pode ser descrito como uma variação do modelo de distribuição normal, idealizado por Gauss, porém com algumas mudanças em relação aos pontos de máximo e mínimo e suas proximidades.

De acordo com Jenkinson (1955), antes dos estudos relacionados aos extremos, as estimativas obtidas para valores máximos e mínimos eram sempre supervalorizados e subvalorizados, respectivamente, dado que as ferramentas utilizadas na época não incluíam no seu sistema de cálculos fenômenos com incidência reduzida.



**Figura 4: Curval de distribuição de Gumbel.**

### **3 - Metodologia**

#### **3.1 – Dados**

Com auxílio de fontes de informações como o terminal da Bloomberg e o software da Thomson Reuters Eikon, foi adquirida uma base de dados dos últimos treze anos com variações diárias nas cotações futuras de contratos de algodão, os preços físicos no dia do fechamento do contrato e atualizações mensais dos últimos dez anos de variáveis divulgadas pelo USDA (United States Department Of Agriculture).

Os dados obtidos são em referência a safra e condições dos Estados Unidos, visto que são o país com maior produção de grãos, grandes exportadores e consumidores e por as bolsas que regularizam o mercado mundial de grãos (Chicago) e as chamadas softs commodities (Nova Iorque) se localizarem nele. Com isso as alterações por ele sofridas afetam de forma mais rápida e de maneira mais contundente a variação dos preços do mercado futuro.

Para se trabalhar em uma escala mensal com todas as variáveis, as cotações diárias são convertidas em valores mensais utilizando-se média móvel simples. Este resultado será utilizado como base para os futuros estudos estatísticos e comparações.

Após a transformação dos dados diários em mensais, os números são divididos em quartis ou quintis para melhor se observar a distribuição dos valores, gerar um histórico e tentar verificar um padrão, ou uma tendência.

### **3.1.1 – Dados dos Preços**

As cotações foram coletadas de janeiro de 2001 a abril de 2013, sendo o valor considerado, a última negociação no fechamento do pregão em Chicago, no caso dos grãos e Nova Iorque, no caso do algodão (valores diários que foram transformados para valores mensais).

Como os valores são baseados em contratos futuros, com datas de vencimentos pré-determinados pelas bolsas, considera-se os valores pagos antecipadamente, como previsões, realizadas através de diferentes métodos, sendo estatísticos ou simplesmente no sentimento dos participantes.

Para os preços reais no encerramento do contrato existem algumas particularidades por se tratar especificamente de commodities, como as datas de vencimento dos contratos. Para o caso do algodão, foi pré-definido pela bolsa de Nova Iorque que os meses nos quais os contratos são fechados são março, maio, julho, outubro e dezembro, assim sendo os valores obtidos são em referência a esses meses. Nos meses nos quais não existem contratos, os valores considerados

são o do próximo vencimento, com exceção do mês de outubro, que por se tratar de um vencimento com pouca liquidez (baixa movimentação), e pouca importância, foi desconsiderado dos dados e a sua referência foi lançada para dezembro.

Ano	Mês de Referência	Previsto	Vencimento	Número do Quartil
<b>2000/01</b>	<b>Jan</b>	55,07	60,38	2,00
	<b>Feb</b>	56,79	60,38	2,00
	<b>Mar</b>	59,68	60,38	3,00
	<b>Apr</b>	56,47	60,85	2,00
	<b>May</b>	57,38	60,85	2,00
	<b>Jun</b>	56,10	56,99	2,00
	<b>Jul</b>	59,86	56,99	3,00
	<b>Aug</b>	64,30	64,93	3,00
	<b>Sep</b>	64,43	64,93	3,00
	<b>Oct</b>	63,35	64,93	3,00
	<b>Nov</b>	64,62	64,93	3,00
	<b>Dec</b>	65,56	64,93	3,00
<b>2001/02</b>	<b>Jan</b>	60,54	50,48	3,00
	<b>Feb</b>	57,70	50,48	3,00
	<b>Mar</b>	50,70	50,48	2,00
	<b>Apr</b>	45,37	43,75	1,00
	<b>May</b>	44,22	43,75	1,00
	<b>Jun</b>	41,01	40,81	1,00
	<b>Jul</b>	41,94	40,81	1,00
	<b>Aug</b>	40,51	36,41	1,00
	<b>Sep</b>	37,08	36,41	1,00
	<b>Oct</b>	31,39	36,41	1,00
	<b>Nov</b>	33,53	36,41	1,00
	<b>Dec</b>	36,60	36,41	1,00

**Figura 5: Tabela dos preços**

### 3.1.2 – Dados das Variáveis

Os dados das variáveis são divulgados pelo USDA todos os meses. Para fins de análise foi pegado o mesmo período em relação aos dados dos preços com o intuito de verificar a significância de todas as variáveis em relação à alteração das cotações. As informações a serem analisadas são área plantada, área colhida, produtividade por área colhida, estoques iniciais, produção, oferta total, uso

doméstico, exportações, uso total (uso doméstico somado com as exportações) e estoques finais.

Como as previsões são feitas com um ano de antecedência, existem valores repetidos que se mantêm inalterados durante alguns meses, dessa forma, na separação dos quintis, os grupos ficaram com quantidade de valores diferentes.

### **3.2 – Análises de dados**

No caso dos preços, foi realizada uma análise para visualizar as diferenças entre o valor previsto com o valor real no vencimento, através de um gráfico com a média da diferença entre os valores dos quartis e um teste-t de diferença entre as médias para identificar se em algum dos quartis, o erro da previsão é mais significativo em relação aos demais e também para verificar a acurácia dos métodos preditivos. Para realizar o método estatístico, foi utilizado o software Gretl.

Para identificar qual das variáveis mais influencia na alteração dos preços, os dados foram sujeitos a uma Análise de Variância (ANOVA), que através de comparações e diferenças entre as médias e as variações, mostra qual o nível de significância entre os parâmetros analisados. Para o teste, os dados foram separados em quintis, sendo que o terceiro quintil, por conter os valores centrais da amostra foi mantido como padrão de análise, uma espécie de grupo de controle. O software STATA foi utilizado para rodar uma ANOVA.

## 4 – Resultados

O primeiro teste realizado com a diferença dos preços (valor previsto – valor no vencimento), através dos gráficos das médias das diferenças dos quartis, indicava que os valores do quarto quartil possuíam os maiores erros de previsão. Assim, de acordo com a teoria comportamental do excesso de confiança, acreditava-se que os produtores, até mesmo nas altas das cotações, sempre desejavam alcançar valores de venda acima do patamar considerado real pelo mercado, pela lei da oferta e da demanda, e também de acordo com os modelos de previsão que possuem como base a curva normal de Gauss, era esperado que nos quartis extremos ocorresse uma diferença significativamente maior em relação aos demais quartis, sendo que no primeiro supunha-se uma subvalorização do preço nas previsões e no quarto uma supervalorização.

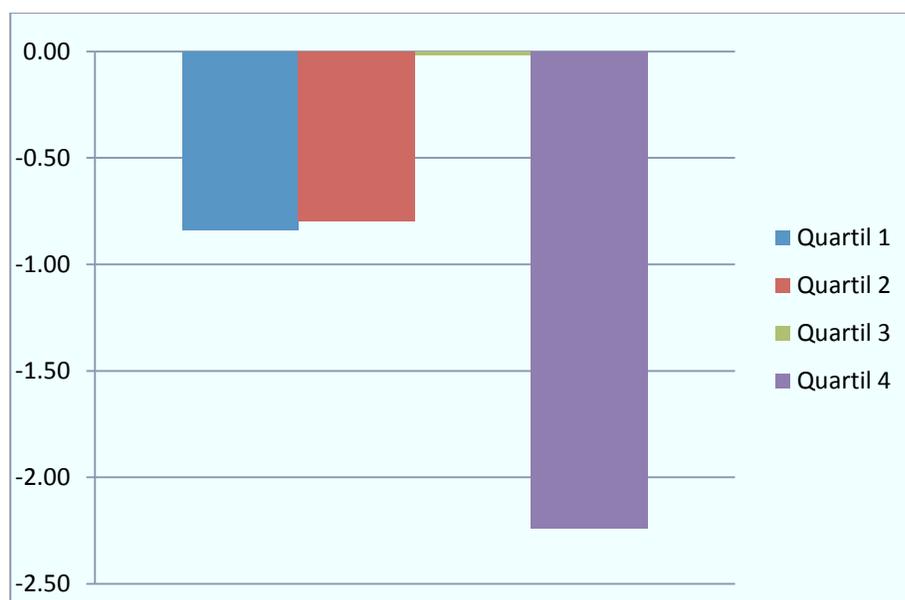
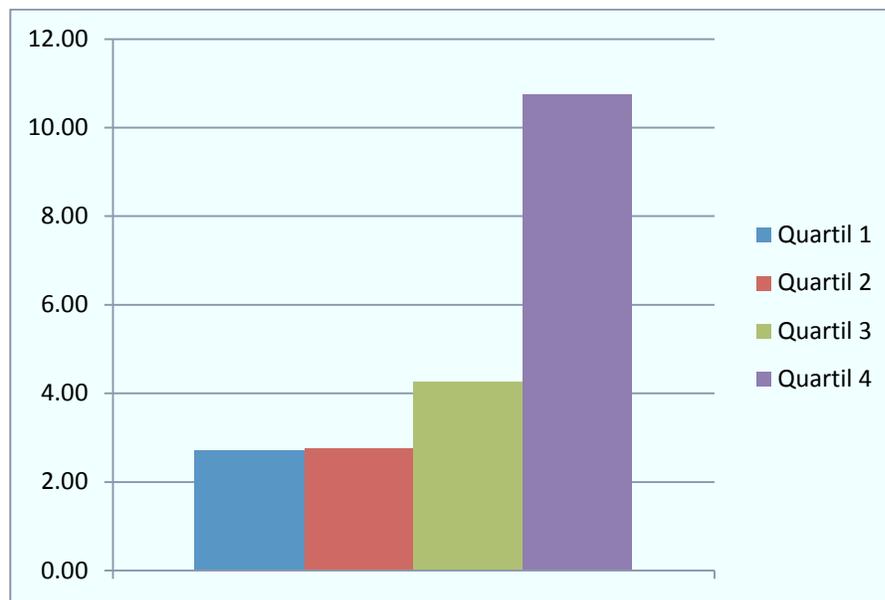


Figura 6: Gráfico da média da diferença de preços: Previsto - Vencimento



**Figura 7: Gráfico da média da diferença de preços em módulo: Previsto - Vencimento**

Através da visualização dos gráficos, a diferença da média entre os quartis é nitidamente maior no quarto quartil. Entretanto, após análises estatísticas mais detalhadas, as suposições iniciais acabam sendo derrubadas, pois com o teste t de Student para diferença de médias, comparando cada previsão do quartil com seu valor real, nota-se que mesmo a diferença dos extremos não é considerada significativa para assumir que as previsões estão desalinhadas com o aumento ou diminuição do seu valor.

Abaixo as respostas geradas pelo software Gretl, cada uma das quatro análises possui duas amostras, a primeira contendo o valor da previsão e a segunda contendo o valor real no vencimento do contrato. Cada amostra foi sujeitada a diferentes operações como cálculo da média, desvio padrão, o erro padrão da média e logo em seguida comparada uma com a outra através do teste t para diferença de médias, sendo que as hipóteses seriam aceitas se valor do teste t, o p-valor, fosse maior do que 0,05 (5%) e rejeitadas se menor.

Primeiro Quartil:

Hipótese nula: Diferença de médias = 0

Amostra 1:  $n = 40$ , média = 43,7055, desvio padrão = 5,12818

Erro padrão da média = 0,810836

Intervalo de confiança de 95% para a média: 42,0654 a 45,3456

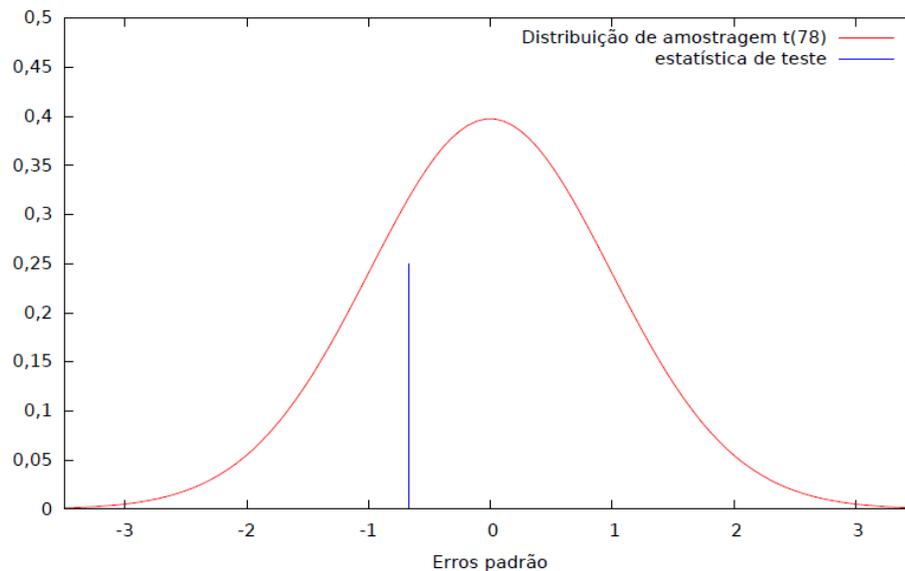
Amostra 2:  $n = 40$ , média = 44,5427, desvio padrão = 5,98244

Erro padrão da média = 0,945906

Intervalo de confiança de 95% para a média: 42,6294 a 46,456

Estatística de teste:  $t(78) = (43,7055 - 44,5427)/1,24587 = -0,671991$

p-valor bilateral = 0,5036



**Figura 8: Gráfico da distribuição do primeiro quartil**

Com os resultados obtidos pelo software, pode-se observar que os erros detectados no primeiro quartil não são considerados significativos a ponto de se afirmar que o problema das análises preditivas está nos métodos utilizados, pois o erro padrão das previsões está elevado e p-valor obtido pelo test-t, está perto dos 50%, muito acima dos 5% necessários para se destacar a hipótese.

## Segundo Quartil

Hipótese nula: Diferença de médias = 0

Amostra 1:  $n = 40$ , média = 53,7316, desvio padrão = 2,18814

Erro padrão da média = 0,345975

Intervalo de confiança de 95% para a média: 53,0318 a 54,4314

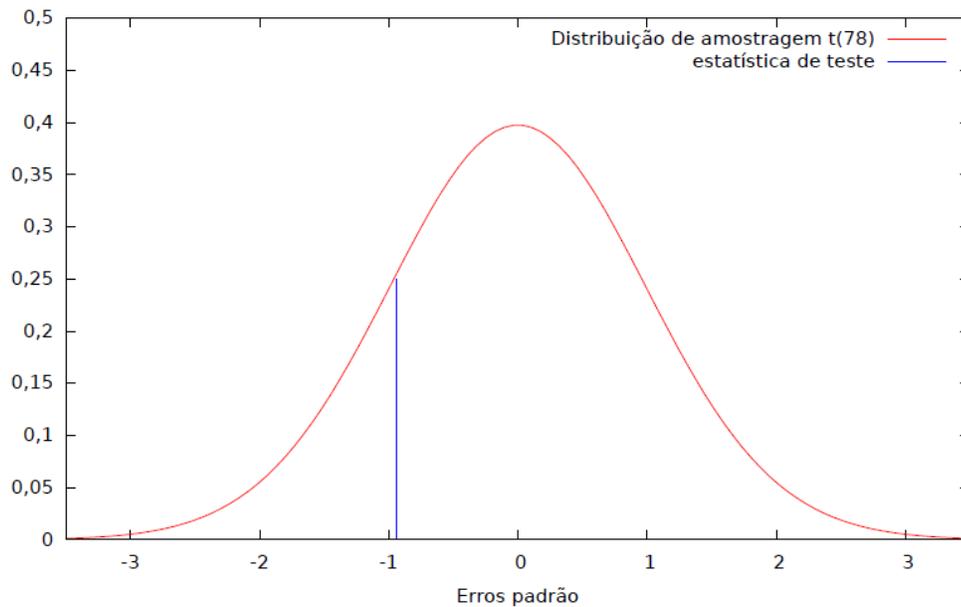
Amostra 2:  $n = 40$ , média = 54,5273, desvio padrão = 4,86092

Erro padrão da média = 0,768579

Intervalo de confiança de 95% para a média: 52,9727 a 56,0819

Estatística de teste:  $t(78) = (53,7316 - 54,5273)/0,842859 = -0,944135$

p-valor bilateral = 0,348



**Figura9: Gráfico da distribuição do segundo quartil**

No segundo quartil, o valor do teste de Student foi o mais baixo entre todos os quartis 34,8%, algo que não era esperado dado às pequenas variações nas diferenças de médias, observados nas figuras 6 e 7 dos gráficos das diferenças das médias, mesmo assim, não se pode afirmar que os erros das previsões são considerados

significativos e os métodos falhos, pois o valor para se negar as hipóteses e dizer que os erros são significativos deve ser abaixo de 5%.

Terceiro Quartil:

Hipótese nula: Diferença de médias = 0

Amostra 1:  $n = 40$ , média = 65,663, desvio padrão = 4,64364

Erro padrão da média = 0,734224

Intervalo de confiança de 95% para a média: 64,1779 a 67,1481

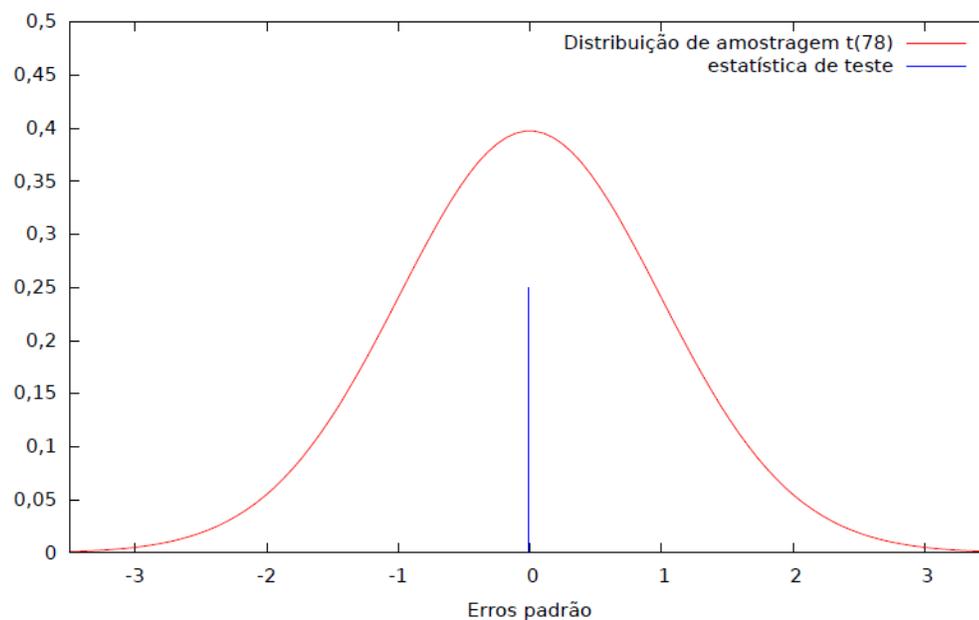
Amostra 2:  $n = 40$ , média = 65,6797, desvio padrão = 8,3982

Erro padrão da média = 1,32787

Intervalo de confiança de 95% para a média: 62,9938 a 68,3656

Estatística de teste:  $t(78) = (65,663 - 65,6797) / 1,51734 = -0,0109939$

p-valor bilateral = 0,9913



**Figura 10: Gráfico da distribuição do terceiro quartil**

Como era esperado para o terceiro quartil, que é o quarto no qual os valores centrais da série estão localizados, os testes mostraram que o p-valor é muito alto,

praticamente 100% (99,13%) o que mostra que os erros das previsões são menores e menos significativos quando os valores da cotação de algodão estão situados nesse intervalo de preço.

Quarto Quartil:

Hipótese nula: Diferença de médias = 0

Amostra 1:  $n = 40$ , média = 101,562, desvio padrão = 34,8542

Erro padrão da média = 5,51094

Intervalo de confiança de 95% para a média: 90,4154 a 112,709

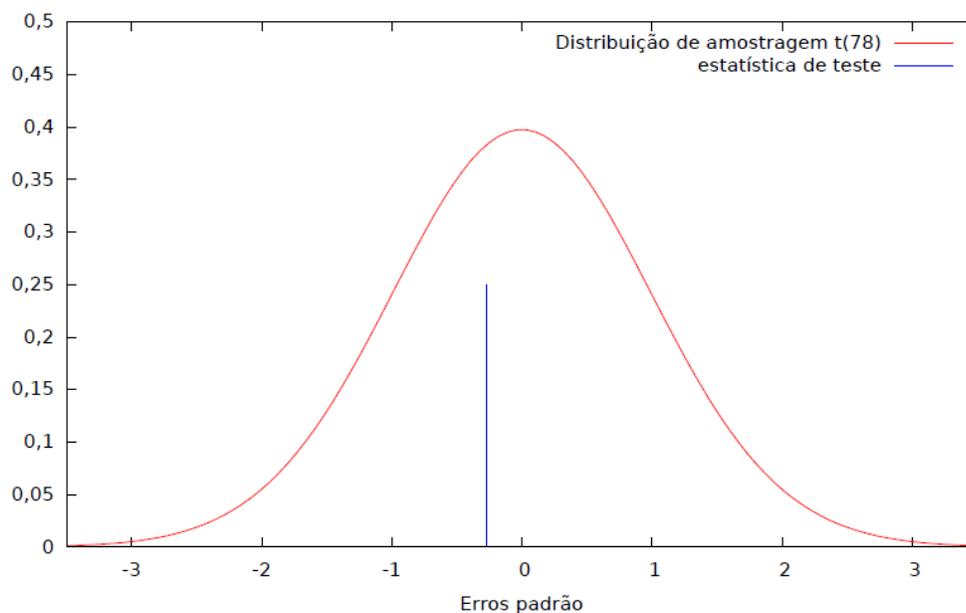
Amostra 2:  $n = 40$ , média = 103,799, desvio padrão = 38,2595

Erro padrão da média = 6,04936

Intervalo de confiança de 95% para a média: 91,5629 a 116,035

Estatística de teste:  $t(78) = (101,562 - 103,799)/8,18323 = -0,273313$

p-valor bilateral = 0,7853



**Figura 11: Gráfico da distribuição do quarto quartil**

No quarto quartil, onde se esperavam erros maiores e mais significativos, o teste-t revelou que os pensamentos iniciais estavam equivocados, pois mesmo contento as maiores diferenças entre os valores previstos e os valores reais, o p-valor se mostrou elevado, na faixa de 78,53%, o que indica que os erros estão coerentes com os quartis anteriores e o método preditivo não supervaloriza as previsões quando os valores do preço do algodão em pluma estão nessa faixa de preço.

Como as hipóteses de todos os testes foram aceitas, pode-se afirmar que os erros em todos os quartis são considerados não significativos, não havendo grande distinção como se havia suposto no início dos estudos. O último quarto representa os maiores erros, entretanto é onde estão localizados os maiores valores da série, sendo assim, percentualmente os erros se mantêm equilibrados ao longo da amostra, o que mostra que os métodos utilizados para se prever os valores futuros na bolsa não estão subvalorizando os valores no primeiro quartil e nem supervalorizado os valores no último quartil, visto que a grande maioria está fundamentada em uma distribuição gaussiana normal que tende a focar na parte central da distribuição.

Para o estudo da influencia das variáveis, o software STATA nos forneceu o quadro referente à ANOVA abaixo:

Number of Obs. = 161                      R-squared = 0.3667  
 Root MSE = 11.2456                      Adj R-squared = 0.155

Source	Partial SS	df	MS	F	Prob > F
Model	8785.593	40	219.6398	1.74	0.0118
AP1	46.74658	1	46.74658	0.37	0.5443
AP2	37.18899	1	37.18899	0.29	0.5886
AP4	1.751451	1	1.751451	0.01	0.9065
AP5	75.04318	1	75.04318	0.59	0.4426
AC1	339.9739	1	339.9739	2.69	0.1037
AC2	13.93439	1	13.93439	0.11	0.7405
AC4	1.838069	1	1.838069	0.01	0.9042
AC5	96.83032	1	96.83032	0.77	0.3833
PR1	104.209	1	104.209	0.82	0.3658
PR2	156.7101	1	156.7101	1.24	0.2679
PR4	6.72113	1	6.72113	0.05	0.8181
PR5	101.6306	1	101.6306	0.8	0.3718
EI1	45.8179	1	45.8179	0.36	0.5484

EI2	405.1186	1	405.1186	3.2	0.076
EI4	22.08726	1	22.08726	0.17	0.6768
EI5	19.41409	1	19.41409	0.15	0.6959
P1	155.3652	1	155.3652	1.23	0.2699
P2	44.56082	1	44.56082	0.35	0.5539
P4	1731.962	1	1731.962	13.7	0.0003
P5	566.7748	1	566.7748	4.48	0.0363
Disp1	24.6349	1	24.6349	0.19	0.6597
Disp2	381.8862	1	381.8862	3.02	0.0848
Disp4	1.397609	1	1.397609	0.01	0.9165
Disp5	15.26732	1	15.26732	0.12	0.7289
CD1	6.298167	1	6.298167	0.05	0.8238
CD2	3.140468	1	3.140468	0.02	0.875
CD4	187.3397	1	187.3397	1.48	0.2259
CD5	185.5108	1	185.5108	1.47	0.2282
Exp1	84.15546	1	84.15546	0.67	0.4163
Exp2	193.3481	1	193.3481	1.53	0.2187
Exp4	3.800834	1	3.800834	0.03	0.8627
Exp5	86.33373	1	86.33373	0.68	0.4103
Uso1	16.50276	1	16.50276	0.13	0.7186
Uso2	0.006307	1	0.006307	0	0.9944
Uso4	155.5701	1	155.5701	1.23	0.2696
Uso5	11.10409	1	11.10409	0.09	0.7675
EF1	111.5764	1	111.5764	0.88	0.3495
EF2	59.02913	1	59.02913	0.47	0.4958
EF4	54.4445	1	54.4445	0.43	0.513
EF5	160.8484	1	160.8484	1.27	0.2617
Residual	15175.64	120	126.4637		
Total	23961.23	160	149.7577		

**Figura 11: Tabela ANOVA gerada pelo software STATA**

Para verificar o grau de significância entre as variáveis e o preço, deve-se observar a coluna Prob > F, quando seu valor for inferior a 0,1 (10%), indica uma significância moderada, mas começa a ser relevante para análises, entretanto quando o valor for menor que 0,05 (5%), significa que existe um alto grau de compatibilidade entre as variações das amostras relacionadas. No caso das variáveis AC1 (área colhida no primeiro quintil) e EI2 (estoques iniciais no segundo quintil), que apresentam valores muito diferentes dos demais quintis da sua amostra, é muito provável que existam erros amostrais, ou a base de dados não foi suficientemente grande para gerar informações mais conclusivas.

Pelo quadro da ANOVA gerado, é possível observar a significância de cada quintil das variáveis e concluir quais são as previsões que mais favorecem a alteração das cotações do algodão em pluma. Dentre as informações divulgadas, as variáveis que apresentaram alto grau de compatibilidade com alteração do preço foram em relação ao quarto e ao último quintil da produção norte-americana, ou seja, quando a produção ultrapassa a barreira do terceiro quintil e avança ao quarto, é quando se observa maiores quantidades de variação no preço do algodão na bolsa de Nova Iorque.

## 5 – Conclusão

Após os estudos realizados, nota-se que apesar da grande evolução e do grande desenvolvimento em relação aos modelos matemáticos de previsão de preço, até mesmo com o auxílio de ferramentas computacionais com inteligência artificial ainda não existem métodos com alta capacidade de acerto.

Mesmo os métodos de redes neurais que simulam o pensamento humano ou os algoritmos genéticos que possuem a capacidade de se desenvolver em busca de melhores resultados ainda estão longe de serem perfeitos e apresentam erros e limitações.

Uma das grandes causas que dificultam a predição com alto nível de acerto se deve ao fato de seres humanos estarem no comando dos negócios e existir a questão da financia comportamental, o sentimento das pessoas ainda influenciam na tomada de decisões se tornando uma variável de importante análise ainda sem condições de serem previstas os mapeadas, pois cada indivíduo possui uma ideia e segue uma linha de raciocínio diferente.

No caso do agronegócio, outro fator que influencia diretamente na produção de uma região e conseqüentemente na previsão do preço do produto e ainda não é possível se realizar uma previsão em longo prazo que se assemelhe a realidade e muito menos de se controlar por meio de tecnologia ou qualquer outro meio existente, é o fator clima. Sem essa variável é praticamente impossível de se prever a produção de uma região, e sendo a produção a variável que mais afeta a cotação do produto, digamos que a previsão para matérias-primas oriundas da agropecuária ainda não podem ser prevista com a clareza necessária para se realizar um planejamento de longo prazo em cima dos dados gerados pelos modelos preditivos.

Pelos dados obtidos é provável que os métodos atuais estejam se sobrecarregando, rodando modelos com variáveis em excesso, fato que contribui com a alta quantidade de erros verificados nas previsões, sendo que como ANOVA mostrou, a variável mais importante para que essa análise seja realizada está na produção norte americana.

## 6 - Referencia Bibliográfica

AGUIAR, D. R. D Mercados Futuros e a Gestão do Risco nos Sistemas Agroindustriais Brasileiro. In: Workshop Brasileiro de Gestão de Sistemas Agroalimentares, 2., 1999. Anais... Ribeirão Preto: USP/Ribeirão Preto, 1999. p.129-136.

ANDERSON, T.W, et Introduction to Multivariate Statistical Analysis, 3. Ed., Nova Iorque, 2003

BM&F. Mercados Futuros: Conceitos e Definições e Mercado de Opções: Conceitos e Definições. Curso de derivativos e opções.

Box, G. E. P., JENKINS, G. M., e REINSEL, G. C. Time Séries Analysis Forecasting and Control. 3rd Ed, Prentice-Hall, EUA, 1994.

BRESSAN, AURELIANO A., Tomada de decisão em futuros agropecuários com modelos de previsão de séries temporais, RAE-eletrônica, v. 3, n. 1, Art. 9, jan./jun. 2004

CEPA (CENTRO DE SOCIOECONOMIA E PLANEJAMENTO AGRÍCOLA). Disponível em: <<http://cepa.epagri.sc.gov.br>>. Acesso em: 5 de maio de 2013

CONAB (COMPANHIA NACIONAL DE ABASTECIMENTO). Disponível em: <<http://www.conab.gov.br>>. Acesso em: 5 maio de 2013.

COPELAND, Thomas E.; WESTON, J. Fred. Financial Theory and Corporate Policy. 3. Ed. EUA: Addison Wiley Publishing Company, 1988

De Bondt,W. F. M. & Thaler, R. H. (1995). Financial decision-making in markets and firms: A behavioral perspective. In Handbooks in Operations Research and Management Science: Finance. Elsevier, Amsterdam. Jarrow, R. A., Maksimovic, R. & Ziemba,W. T. (orgs.).

DOS SANTOS, ELISANDRA; WOLFF, LAION; SOUZA, ADRIANO M., Transmissão e a influência do volume dos estoques públicos sobre o preço do arroz no Brasil, Revista Ciência Rural, Santa Maria – RS, 2013

EGELKRAUT, T.; B. SHE RRICK, P. GARCIA; PENNING, J. 2006. Producers' yield and yield risk: perceptions vs. reality. Long Beach: American Agricultural Economics

Association, 2006. Disponível em: <<http://ageconsearch.umn.edu/handle/21369>>. Acesso em: 6 maio de 2013.

EVANS, J. Psychological Pitfalls in Forecasting. *Futures*. v. 14, n.4, p.258-265, 1982

FERREIRA, D. F. Análise multivariada. Lavras, 1996.

HAIR, J. F.; ANDERSON, et al. Análise multivariada de dados. 6. ed. Porto Alegre, 2009.

JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quartely Journal of the Royal Meteorological Society*, v.81, p158-171, Apr. 1955.

McCULLOCH, W., PITTS, W., A logical calculus of the ideas immanent in neurons activity, *Bulletin of Mathematical Biophysics* 5, 1943.

PINDYCK, R. S.; RUBENFIELD, D. L. *Econometric Models and Economic Forecasts*. New York, McGrawHill, 3rd Ed., 1991. 565p.

TURBAN, E. *Decision Support and Expert Systems: Management Support Systems*. New York, MacMillan, 1993. 832p.

VICINI, LORENA – Análise multivariada da teoria a prática, Santa Maria –RS, 2006

VERE, D.T.; GRIFFITH, G. R. Comparative Forecast Accuracy in the New South Wales Prime Lamb Market. in.: *Australian Journal of Agricultural Economics*, vol. 34, n°2, p.103-117, 1990

Zirilli, Joseph S., *Financial Prediction using Neural Networks* International Thomson Publishing, London, 1997