

Modelo inteligente para previsão do preço diário da saca do café baseado em combinação de previsores

Leandra de Carvalho Nogueira¹

Ricardo Menezes Salgado²

Denismar Alves Nogueira³

Resumo: A produção, comercialização e mercado do café geram renda para milhares de pessoas em todo o mundo, além de ser uma cultura tradicional no Brasil e um importante segmento da agricultura brasileira. O mercado de produtos agrícolas é caracterizado por possuir maior sensibilidade na oscilação dos preços. Um dos objetivos das previsões econômicas é a redução de incerteza, um fator de extrema importância no setor agropecuário, que sofre constantes distúrbios irregulares no valor de seus produtos. Considerando estes fatores, pode-se perceber que a produção cafeeira necessita de instrumentos que auxiliem a tomada de decisão dos agentes envolvidos neste agronegócio. Nesta ótica, este trabalho propõe a elaboração de uma metodologia de previsão do preço do café à curto prazo, baseada na combinação de vários previsores distintos. Os resultados mostraram que a combinação dos previsores apresentou, na maioria dos casos, resultados mais acurados quando comparados aos obtidos pelas componentes de previsão individualmente.

Palavras-chave: *Previsão de preço do café, Combinação de previsores, Séries temporais.*

1 Introdução

O café arábica é uma cultura tradicional do Brasil, sendo um importante produto da agricultura brasileira. Segundo o SindiCafé (Sindicato da Indústria de Café do Estado de Minas Gerais, 2004), ele é um dos mais valiosos produtos primários comercializados no mundo, sendo superado apenas em valor pelo petróleo, como origem do desenvolvimento de negócios entre os países. Seu cultivo, processamento, comercialização, transporte e mercado proporcionam milhões de empregos em todo o mundo.

O mercado de produtos agrícolas caracteriza-se por apresentar maior grau de sensibilidade no que se refere a oscilações de preços. Isso é decorrência direta das próprias características intrínsecas que

¹ UNIFAL-MG - Universidade Federal de Alfenas. Email: leandra.carvalho@bcc.unifal-mg.edu.br

² UNIFAL-MG - Universidade Federal de Alfenas. Email: ricardo@bcc.unifal-mg.edu.br

³ UNIFAL-MG - Universidade Federal de Alfenas. Email: denisnog@gmail.com

regem as condições de produção do mercado agrícola, que não somente proporcionam elevado grau de instabilidade, como também grande amplitude de variação dos preços de seus produtos. (Santiago *et al.*, 1996).

Um dos objetivos das previsões econômicas é a redução de incerteza, que é de suma importância no setor agropecuário, que está sujeito a constantes distúrbios irregulares.

Neste contexto, pode-se perceber que a produção cafeeira necessita de instrumentos que auxiliem na tomada de decisão de produtores, compradores e investidores em geral que participam deste agronegócio. A previsão diária do preço do café pode auxiliar a tomada de decisão destes agentes, servindo como base para se escolher o melhor momento para vender, comprar ou estocar o produto.

2 Materiais e Métodos

Segundo Leone Filho (2006), a utilização de múltiplos previsores tenta explorar o bom comportamento local de cada um dos previsores e, com isto, aumentar a precisão e a confiabilidade da previsão, uma vez que se um dos previsores errar muito em determinado subconjunto de dados de entrada, os outros previsores tendem a compensar o erro.

A ideia principal do modelo proposto consiste em realizar a previsão diária do preço do café e combinar os resultados. Para cada mês previsto, foram utilizados os dez últimos anos para a realização dos experimentos, considerando apenas os meses equivalentes ao almejado. A escolha do número de anos anteriores utilizados foi feita através de busca exaustiva. Desta forma, a previsão foi realizada dia a dia para cada mês considerado. Neste contexto, para cada mês almejado, foram selecionadas as duas componentes que obtiveram os menores índices de erro de previsão. Uma vez selecionadas, os resultados são combinados e uma análise comparativa entre os erros de previsão é efetuada.

2.1 Componentes de Previsão

O ajuste de cada componente foi feito utilizando busca exaustiva para selecionar os parâmetros. Estas componentes geradas realizam a previsão dia a dia para um determinado mês. O desempenho das componentes será avaliado através do erro percentual absoluto médio (MAPE), calculado da seguinte forma: $MAPE = 100 \frac{1}{n} \sum_{i=1}^n \left| \frac{X_{previsto} - X_{real}}{X_{previsto}} \right|$, onde $X_{previsto}$ representa o valor previsto e X_{real} representa o valor real. A seleção das componentes ocorre em função daquelas que obtiveram o menor MAPE.

2.2 Combinadores

Neste trabalho, foram utilizadas duas formas de combinar os resultados das componentes: média aritmética e média ponderada, calculadas das seguintes formas:

A primeira forma consiste no uso de média aritmética, $MA = \frac{(V_1+V_2+\dots+V_n)}{n}$, sobre os resultados das melhores componentes, selecionadas a partir do menor MAPE, onde V_n representa o valor do preço do café previsto para cada uma das componentes e n representa o número total de componentes utilizadas.

A segunda maneira de se construir o combinador consiste no uso da média ponderada, $MP(p_1, p_2, \dots, p_n) = \frac{(p_1V_1+p_2V_2+\dots+p_nV_n)}{\sum_{i=1}^n p_i}$, sobre os resultados das melhores componentes selecionadas. Nesta equação, p representa os pesos, distribuídos de forma que as componentes que alcançaram menores índices de erros recebem os maiores valores no processo e V_n denota os valores previstos.

3 Resultados e Discussões

Para testar a eficiência do modelo proposto foram realizadas simulações utilizando diferentes anos e diferentes períodos destes anos. O modelo de previsão foi testado em onze períodos distintos: Março/2008, Junho/2008, Dezembro/2008, Março/2010, Junho/2010, Dezembro/2012, Março/2012, Junho/2012, Dezembro/2012, Janeiro/2013 e Março/2013.

3.1 Resultados das Componentes Individuais

A partir da análise da Tabela 1, é possível perceber que nem sempre as melhores componentes selecionadas para a fase de validação se comportam da mesma maneira para a fase de teste. Para exemplificar, ao observar o mês de dezembro de 2008, a componente que alcançou o melhor resultado de previsão na validação teve o mesmo comportamento da segunda componente, que obteve um erro maior na fase da validação.

| Tabela 1: Erro Percentual Absoluto Médio Verificado Para as Componentes Selecionadas | | | | |
|---|----------------------|-------------------|----------------|------------|
| | Modelo da Componente | Número de Atrasos | Erro Validação | Erro Teste |
| Março de 2008 | GRAD | 5 | 3,64% | 4,52% |
| | FR | 6 | 3,68% | 4,71% |
| Junho de 2008 | OSS | 6 | 1,18% | 1,48% |
| | DFP | 5 | 1,19% | 1,56% |
| Dezembro de 2008 | BPM | 7 | 0,97% | 1,13% |
| | GRAD | 6 | 1,01% | 1,13% |
| Março de 2010 | PR | 3 | 0,61% | 0,72% |
| | PR | 2 | 0,62% | 0,76% |
| Junho de 2010 | BPM | 6 | 1,35% | 1,65% |
| | BPM | 4 | 1,76% | 1,80% |

| | | | | |
|------------------|------|---|-------|-------|
| Dezembro de 2010 | FR | 7 | 2,26% | 2,74% |
| | PR | 2 | 2,27% | 3,05% |
| Março de 2012 | BPM | 5 | 1,49% | 2,00% |
| | BPM | 7 | 1,77% | 2,02% |
| Junho de 2012 | SCGM | 2 | 4,96% | 4,22% |
| | SCGM | 6 | 4,97% | 4,24% |
| Dezembro de 2012 | PR | 6 | 0,59% | 1,53% |
| | PR | 4 | 0,72% | 1,67% |
| Janeiro de 2013 | OSS | 3 | 1,07% | 1,27% |
| | FR | 2 | 1,09% | 1,30% |
| Março de 2013 | OSS | 5 | 1,52% | 2,00% |
| | PR | 7 | 1,63% | 2,18% |

Vale destacar que, eventualmente pode acontecer de uma componente selecionada na fase de validação apresentar um melhor desempenho na fase de teste, conforme ocorrido no mês de Junho de 2012.

Um acontecimento considerado ideal e desejável seria que os erros obtidos pelas componentes selecionadas fossem numericamente menores ou iguais aos erros de previsão alcançados na fase de validação, por essas mesmas componentes. De uma forma geral, as componentes selecionadas conseguiram alcançar um desempenho interessante nos dias de teste.

3.2 Resultados dos Combinadores

A Tabela 2 foi estruturada de modo que foram abrangidos os valores de MAPE dos previsores combinados, facilitando a análise dos resultados.

| Tabela 2: Erro Percentual Absoluto Médio Verificado Para os Combinadores | | | | | | | | | | | |
|--|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| | Mar/2008 | Jun/2008 | Dez/2008 | Mar/2010 | Jun/2010 | Dez/2010 | Mar/2012 | Jun/2012 | Dez/2012 | Jan/2013 | Mar/2013 |
| MA | 4,51% | 1,50% | 1,12% | 0,74% | 1,68% | 2,89% | 1,96% | 4,23% | 1,60% | 1,27% | 2,09% |
| MP (6,4) | 4,48% | 1,50% | 1,12% | 0,73% | 1,66% | 2,86% | 1,96% | 4,23% | 1,59% | 1,28% | 2,07% |
| MP (7,3) | 4,46% | 1,49% | 1,11% | 0,73% | 1,65% | 2,83% | 1,97% | 4,23% | 1,57% | 1,28% | 2,05% |
| MP (8,2) | 4,47% | 1,49% | 1,11% | 0,72% | 1,65% | 2,80% | 1,98% | 4,22% | 1,56% | 1,29% | 2,04% |
| MP (9,1) | 4,50% | 1,49% | 1,12% | 0,72% | 1,65% | 2,77% | 1,99% | 4,22% | 1,54% | 1,30% | 2,02% |
| Melhor | 4,52% | 1,48% | 1,13% | 0,72% | 1,65% | 2,74% | 2,00% | 4,22% | 1,53% | 1,27% | 2,00% |
| Pior | 4,71% | 1,56% | 1,13% | 0,76% | 1,80% | 3,05% | 2,02% | 4,24% | 1,67% | 1,30% | 2,18% |

Avaliando a Tabela 2 verifica-se que, para cada mês, o combinador pode ser mais eficiente que outros ou até mesmo nenhum deles se sobressair, perdendo algum desempenho, em termos de precisão

no resultado, para as componentes individuais, como ocorreu nos meses de junho de 2008, dezembro de 2010, dezembro de 2012 e março de 2013.

A Tabela 2 é caracterizada por apresentar valores de MAPE que empatam entre diferentes combinadores. Este empate ocorre em sete dos onze meses previstos: junho de 2008, dezembro de 2008, março de 2010, junho de 2010, março de 2012, junho de 2012 e janeiro de 2013. Os combinadores apresentaram resultados melhores que a componente individual para os seguintes meses: março de 2008, dezembro de 2008 e março de 2012. Outra situação que pode ocorrer é o combinador manter o mesmo MAPE que a melhor componente individual. Este cenário ocorreu em março de 2010, junho de 2010, junho de 2012 e janeiro de 2013.

Os combinadores, de uma forma geral, possuem uma boa capacidade de generalização, aumentando a precisão dos resultados em alguns casos e mantendo em outros.

4 Conclusão

A previsão de preços da saca do café não é uma tarefa trivial, pois possui uma grande quantidade de variáveis envolvidas, como as interferências exógenas já citadas, entre outros fatores, que dificultam a previsão de valores à longo prazo.

As análises realizadas demonstraram que os resultados das componentes individuais de previsão, quando submetidos a um determinado tipo de combinação, se comportam de maneira semelhante ou superior a melhor componente individual, em boa parte dos meses considerados na fase de teste, e superando a pior componente individual para todos os casos.

Portanto, a utilização de um combinador de previsão em substituição à utilização de apenas uma componente, tratada de forma individual, pode ser considerada válida e eficaz.

Referências

- [1] LEONE FILHO, M.A.; CAMARGO, M.L.B.; MARGARIDO, M.A. *Ensembles of Selected and Evolved Predictors using Genetic Algorithms for Time Series Prediction*, IEEE World Congress on Computational Intelligence, Vancouver, 2006.
- [2] SANTIAGO, M.M.D.; CAMARGO, M.L.B.; MARGARIDO, M.A. Detecção e Análise de Outliers em Séries Temporais de Índices de Preços Agrícolas no Estado de São Paulo. Agricultura em São Paulo, SP, pp.89-115, 1996.
- [3] SINDICAFÉ – Sindicato da Indústria de Café do Estado de Minas Gerais, Agosto de 2004. Disponível em < <http://sindicafe-mg.com.br/plus/modulos/conteudo/?tac=cafe-no-mundo> >. Acessado em 25 de Fevereiro de 2013.