UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

PRODUTIVIDADE DE CAFÉ ARÁBICA ESTIMADA A PARTIR DE DADOS DE MODELOS DE CIRCULAÇÃO GLOBAL

Taynara Tuany Borges Valeriano

Engenheira Agrônoma

2017

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

PRODUTIVIDADE DE CAFÉ ARÁBICA ESTIMADA A PARTIR DE DADOS DE MODELOS DE CIRCULAÇÃO GLOBAL

Taynara Tuany Borges Valeriano

Orientador: Prof. Dr. Glauco de Souza Rolim

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Produção Vegetal).

2017

Ficha catalográfica

Valeriano, Taynara Tuany Borges
A163p Produtividade de café arábica estimada a partir de dados de modelos de circulação global / Taynara Tuany Borges Valeriano. – – Jaboticabal, 2017
xiii, 88 p. : il. ; 29 cm

Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências Agrárias e Veterinárias, 2017
Orientador: Glauco de Souza Rolim
Banca examinadora: Lucieta Guerreiro Martorano, Márcio José de Santana
Bibliografia

1. Modelagem. 2. *Coffea arabica.* 3. Dados em GRID. 4. *Big data*. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e Veterinárias.

CDU 633.73:551.5

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Diretoria Técnica de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.

unesp

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO: PRODUTIVIDADE DE CAFÉ ARÁBICA ESTIMADA A PARTIR DE DADOS DE MODELOS DE CIRCULAÇÃO GLOBAL

AUTORA: TAYNARA TUANY BORGES VALERIANO
ORIENTADOR: GLAUCO DE SOUZA ROLIM

Aprovada como parte das exigências para obtenção do Título de Mestra em AGRONOMIA (PRODUÇÃO VEGETAL), pela Comissão Examinadora:

Prof. Dr. GLAUCO DE SOUZA ROLIM
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Márcio José de Santana

Prof. Dr. MÁRCIO JOSÉ DE SANTANA
Departamento de Agronomia / IFTM - Uberaba/MG

Pesquisadora Dra. LUCIETA GUERREIRO MARTORANO
Departamento de Agrometeorologia e Georeferenciamento / EMBRAPA - Santarém,PA

Jaboticabal, 21 de fevereiro de 2017.

DADOS CURRICULARES DO AUTOR

TAYNARA TUANY BORGES VALERIANO – nascida em Araxá, Minas Gerais, no dia 04 de julho de 1991, filha de Geraldo Humberto Valeriano (*in memorian*) e Magna Regina Borges, natural de Pratinha - MG. Cursou o Ensino Fundamental na Escola Estadual Marlene Martins Reis, deste município, e Ensino Médio nas Escolas Centro de Educação Ciência e Tecnologia (CEFET) – Câmpus Araxá e Escola Estadual Dom José Gaspar, na cidade de Araxá - MG, tendo finalizado no ano de 2009. Ingressou no Ensino Superior no ano de 2010 no curso de Engenharia Agronômica pelo Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro (IFTM), Câmpus Uberaba, obtendo o título de Engenheira Agrônoma, em fevereiro de 2015. Durante a graduação foi bolsista do Programa de Educação Tutorial (PET) por três anos e meio e voluntária no programa de Iniciação Científica por 4 anos e meio nas áreas de Agrometeorologia, irrigação e fitotecnia, sob orientação do Prof. Dr. Márcio José de Santana. Colaborou na execução e apresentação de trabalhos científicos, com resumos e artigos publicados em periódicos especializados. Em agosto de 2015, iniciou o curso de Mestrado em Agronomia, no Programa de Produção Vegetal, com concentração na área de Modelagem Agrometeorológica, pela Universidade Estadual Paulista “Júlio de Mesquita Filho” – Câmpus de Jaboticabal, São Paulo, no Group of Agrometeorological Studies (GAS), localizado no Departamento de Ciências Exatas, sob orientação do Prof. Dr. Glauco de Souza Rolim. No decorrer do curso realizou pesquisas com foco em modelagem agrometeorológica para o cafeeiro a partir de dados em grid, sendo essa a área de maior concentração dos estudos. Em fevereiro de 2017, submeteu-se à banca examinadora para obtenção do título de Mestre em Agronomia.

“A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original”.

Albert Einsten

Aos meus pais Geraldo Humberto Valeriano (*in memorian*) e Magna Regina Borges que me criaram com todo o amor, carinho e educação que se pode dar a um filho.

A vocês**,**

DEDICO!

A minha irmã, familiares e amigos,

OFEREÇO!

AGRADECIMENTOS

Em primeiro lugar agradeço a Deus por ter sempre me amparado nas minhas escolhas para que não me arrependesse e nem desistisse delas, por mais difíceis que fossem.

À Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências Agrárias e Veterinárias, Campus de Jaboticabal, em especial ao Programa de Pós-Graduação em Agronomia (Produção Vegetal) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão de bolsa e oportunidade de realizar um curso de pós-graduação.

À toda minha família em especial minha mãe, Regina, minha irmã Gabriella, minha avó Anacleta, minhas Tias Márcia e Rita, meus padrinhos Luiz e Joana, meus primos João Antônio e Juliana, meu padrasto Neirimar, meus afilhados, Luís Guilherme, Isadora e Adrian Lucas, minha sobrinha de coração Estela e meu tio Lázaro, pela paciência e pela ausência em suas vidas, durante todos esses anos de estudos até aqui. Obrigado por vocês acreditarem em mim!

À minha prima Stefany Souza, pela cumplicidade, lealdade, carinho e amizade de todos esses anos.

Ao meu orientador Prof. Dr. Glauco de Souza Rolim que tenho uma admiração e consideração imensa pelo seu trabalho, bem como pela referência que é para todos nós, orientandos. Muito obrigado pelos ensinamentos passados e pela confiança em mim depositada.

As minhas irmãs de coração, Rayeny Ávila e Sayeny Ávila as quais sempre estiveram ao meu lado me apoiando e incentivando. Aos amigos de Jaboticabal, Bárbara, Edgard, Adão, Jordana e Larissa e a todos da Rep. Toca Fogo que foram fundamentais para adaptação em Jaboticabal.

Aos meus amigos e parceiros do GAS, José Reinaldo, Victor Moreto, Lucas Aparecido, Kamila Menezes, Juliana, João e Clayson. Obrigada pela convivência e troca de experiência durante esse tempo.

Aos amigos de longa data, Isabela, Ana Cristhina, Kárita, Ana Gabriela, Letícia, Vanessa, Mariana e Cristina que, mesmo de longe torceram por mim e entenderam minha ausência.

Ao Emmanuel Moreira, por todo amparo e carinho.

Aos membros da banca de qualificação, Dr. Diego, Prof. Dr. Rogério e Prof. Dr. Newton, que muito contribuíram para a construção desta dissertação.

Aos membros da banca da defesa, Prof. Dr. Márcio José de Santana e Dr. Lucieta Guerreiro Martorano, que muito contribuíram para a dissertação.

Aos colegas do Departamento de Ciências Exatas, Zezé, Shirlei, Adriana, Mara, Daniel, Bruna, Katherine, Gustavo, Prof. Peruzzi, Prof. Newton, Carlão e Vanessa, muito obrigada pela convivência durante esse tempo.

A todos que de alguma forma contribuíram para que chegasse hoje aqui, o meu muito obrigada!

SUMÁRIO

PRODUTIVIDADE DE CAFÉ ARÁBICA ESTIMADA A PARTIR DE DADOS DE MODELOS DE CIRCULAÇÃO GLOBAL

RESUMO

O Brasil é o maior produtor da segunda *commodity* de maior valor no mercado, o café. O conhecimento de técnicas eficazes de estimativa de produtividade é de grande importância para o mercado cafeeiro, possibilitando melhor planejamento e tornando a atividade mais sustentável. Uma forma eficaz de se estimar a produtividade é através da modelagem agrometeorológica, que quantifica a influência do clima nos cultivos agrícolas. Entretanto, dados climáticos são necessários, e estes em sua maioria são provenientes de estações de superfície. Uma forma de inovar essa técnica é a utilização de outra fonte de dados climáticos, como os dados em grid (GD) que são resultado da combinação de diversas fontes, como sondas oceânicas, sensoriamento remoto, entre outros. O presente trabalho teve como objetivo estimar a produtividade de *Coffea arabica* a partir do modelo proposto por Santos e Camargo (2006) (SC) com dados meteorológicos provenientes dos GD dos sistemas ECMWF e NASA para regiões cafeeiras de Minas Gerais e São Paulo. Em um primeiro momento foram comparados dados de temperatura do ar e precipitação obtidos pelo ECMWF e NASA aos dados de estações meteorológicas de superfície, com o objetivo de verificação da acurácia dos GDs. No segundo momento foi proposta uma calibração no modelo de estimativa de produtividade de Santos e Camargo (2006), para a utilização dos GDs. Para temperatura, os dados do ECMWF e NASA foram precisos e acurados, com valores mínimos de RMSE iguais a 0,37 e 0,50 °C, e d de Willmott iguais a 0,86 e 0,53, respectivamente. Para precipitação os valores mínimos de precisão e acurácia foram inferiores, com RMSE iguais a 2,15 e 5,33 °C, e d de Willmott iguais a 0,79 e 0,59, respectivamente. Quando os dados GD foram aplicados no modelo SC para estimação de produtividade de café de forma geral o ECMWF foi superior a NASA. A calibração dos coeficientes do modelo SC para dados do ECMWF e NASA foi eficiente, pois houve diminuição do erro percentual absoluto médio (MAPE), raiz do erro médio quadrático sistemático (RMSEs) e raiz do erro médio quadrático não-sistemático (RMSEn) em relação aos modelos SCs que usaram dados de superfície. Estes resultados comprovam que os modelos SCs foram, precisos, acurados e apresentaram baixa tendência, portanto a utilização de dados meteorológicos de GD nos modelos SCs foi possível e recomenda-se a utilização dos mesmos para as regiões cafeeiras de Minas Gerais e São Paulo.

PALAVRAS-CHAVES: Modelagem. *Coffea arabica*. Dados em GRID. Big data.

ESTIMATED ARABIC COFFEE YIELD FROM GLOBAL CIRCULATION MODEL DATA

ABSTRACT

Brazil is the largest producer of the second most valuable commodity on the market, coffee. The knowledge of effective techniques of yield estimation is importance for the coffee market, enabling better planning and making the activity more sustainable. An effective way of yield estimating is through agrometeorological modeling, which quantifies the influence of climate on agricultural crops. However, weather data is needed, and these mostly come from surface stations. One way to innovate this technique is to use another source of climate data, such as Gridded data (GD), which are the result of a combination of diverse sources such as oceanic probes, remote sensing, among others. This work aimed to propose the use of the model of yield estimation proposed by Santos and Camargo (2006) with meteorological data from the GCMs, ECMWF and NASA for coffee regions of Minas Gerais and São Paulo. At first, temperature and precipitation data obtained through the ECMWF and NASA were compared and verified with data from surface meteorological stations. In the second moment, a calibration was proposed in Santos and Camargo (2006), for the use of GDs (ECMWF and NASA). For temperature, ECMWF and NASA data were accurate and accurate, with minimum RMSE values of 0.37 and 0.50 ° C, and Willmott's d values of 0.86 and 0.53, respectively. For precipitation, the minimum values of precision and accuracy were lower, with RMSE equal to 2.15 and 5.33 ° C, and d of Willmott equal to 0.79 and 0.59, respectively. When the GD data were applied in the SC model for coffee yield estimation the ECMWF was superior to NASA. The calibration of the SC model coefficients for ECMWF and NASA data was efficient, since there was a decrease in the mean absolute percentage error (MAPE), mean root mean square root (RMSEs) and root mean square root mean square error (RMSEn) in SCs that used surface data.These results confirm that the models were accurate, accurate and presented a low tendency, so the use of GD meteorological data in agrometeorological models was satisfactory and it is recommended to use them for the coffee regions of Minas Gerais and São Paulo.

KEYWORDS: Modeling. *Coffea arabica*. Grid Data. Big data.

CAPÍTULO 1 – Considerações gerais

1 – INTRODUÇÃO

A interação entre as plantas e o ambiente envolve uma complexidade de processos físicos, químicos e biológicos. Visando obter um conhecimento mais profundo sobre as respostas do cultivo ao ambiente, modelos de estimativa de produtividade são utilizados como ferramentas, possibilitando o estudo e o entendimento do sistema agrícola (OLIVEIRA et al., 2011).

A fonte tradicional dos dados climáticos para a entrada em modelos agrometeorológicos são as estações meteorológicas de superfície. Esses dados são limitados pois em várias regiões do mundo existem um número geralmente insuficiente de estações em relação à extensão da região a ser estudada. Outras limitações são relacionadas aos vários possíveis erros relacionados à obstrução de sensores, perda de calibração e problemas no armazenamento de dados (MORAES; ROCHA; LAMPARELLI, 2014).

Uma alternativa viável para solucionar estes problemas é a utilização de dados em grade (GD). Os GDs são sistemas que combinam diversas fontes de dados: dentre sondas oceânicas, balões meteorológicos, sensoriamento remoto e até mesmo as estações de superfície e disponibilizam as estimações e previsões em forma de grade (grid), dividindo o globo terrestre em três dimensões. Alguns GDs que podem ser citados são European Center for Medium-Range Weather Forecast e NASA's Global Modeling and Assimilation Office (NASA/GMAO).

Os dados de GDs tem sido amplamente estudado por autores como Albergel et al. (2012) na Austrália, Holmes et al. (2012) em Oklahoma (USA), Wang; Song ; Cao, (2012) na China, Rubel e Rudolf (2001) nos Alpes Europeus, Jhohann et al. (2011) no Brasil, porém mais estudos são necessários para verificar a qualidade destes dados em relação aos de superfície.

Devido a grande relevância do cultivo do cafeeiro, visto que, o Brasil tem a maior produção do mundo (AMARASINGHE et al., 2015), o desenvolvimento de modelos agrometeorológicos com o objetivo de aferir essa influência dos elementos meteorológicos, torna-se uma questão estratégica e de planejamento para o país, auxiliando no abastecimento interno e relações no mercado externo.

O modelo de estimativa de produtividade para café mais utilizado no Brasil é o desenvolvido por Santos e Camargo (2006). Este modelo tem por base o de quebra relativa proposto por Camargo et al. (2003), que é constituído por três componentes, fenológico, térmico e hídrico. Santos (2005) parametrizou os coeficientes deste modelo e inseriu um quarto componente, que leva em consideração a bienalidade produtiva do cafeeiro, ao inserir na equação uma penalização em relação a produtividade do ano anterior.

Diante do exposto o objetivo neste trabalho foi avaliar os dados provenientes do ECMWF e NASA, e calibrar o modelo de estimativa de produtividade para café, proposto por Santos e Camargo (2006), a partir de dados meteorológicos oriundos dos GD, ECMWF e NASA, em regiões cafeeiras de São Paulo e Minas Gerais.

1.1 REVISÃO DE LITERATURA

1.1.1 Modelagem agrometeorológica

De acordo com o IBGE (2016) a previsão de produtividade de diversas cultivos em todo o Brasil é realizada com uso de informações municipais, obtidas por meio de um sistema de levantamento, baseado em opiniões de agentes técnicos e econômicos relacionados ao setor. Contudo, além de serem passíveis de manipulação, essas informações não permitem uma análise quantitativa dos erros envolvidos, uma vez que estas possuem um caráter subjetivo dos levantamentos.

Atualmente a Companhia Nacional de Abastecimento (CONAB), vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA), é a responsável oficial pelo levantamento e divulgação da previsão da safra cafeeira (CARVALHO et al., 2005). Em Minas Gerais, por exemplo, os técnicos da CONAB visitam municípios produtores de café, onde colhem informações juntos aos órgãos de assistência técnica, cooperativas e entidades ligadas ao setor. Porém, de acordo Carvalho et al. (2005) no Brasil apesar do levantamento das informações ser calculado em estudo estatístico e científico, a determinação da produtividade ainda é subjetiva.

Diante deste contexto, a utilização de modelos agrometeorológicos de estimativa de produtividade tem ganhado destaque no monitoramento das safras, pois é uma importante ferramenta para quantificar o efeito do clima e o crescimento e desenvolvimento da planta (ROLIM et al., 2011), proporcionando o planejamento dos mercados interno e externo.

No Brasil existem vários modelos de estimativa de produtividade para várias cultivos como, soja (SILVA-FUZZO et al., 2015); arroz (KLERING et al., 2008), jiló (ROLIM et al., 2011), amendoim (MORETO; ROLIM, 2015), café (SANTOS; CAMARGO, 2006) entre outros. Um fator limitante desses modelos é a origem dos dados de entrada, sendo que em sua maioria são provenientes de estações meteorológicas de superfície.

De acordo com Pereira et al. (2002), o Brasil não apresenta uma rede de estações que atenda todas suas necessidades além de um maior número se concentrar em áreas mais desenvolvidas do que em outras mais remotas. Outro fator limitante é a presença de dados faltantes nas séries climáticas que acaba limitando diversas pesquisas, dificultando conclusões significativas sobre a gestão de recursos naturais (MOELETSI; WALKER, 2012). Uma forma de solucionar este problema é a utilização de dados meteorológicos em grid.

1.1.2 Dados em Grid

A utilização de dados meteorológicos em grid (GD) tem ganhado destaque no monitoramento climático em cultivos agrícolas nos últimos anos (GASPAR et al., 2015). Os GDs são modelos que combinam informações de diversas fontes, superfície, oceanos, sensoriamento remoto e fazem estimações e previsões das condições meteorológicas separando a superfície e a atmosfera do planeta em um grid de 3 dimensões (Bechtold et al., 2008).

Alguns GDs que podem ser citados são modelo atmosférico global do INPE/ CPTEC, o Sistema Global de Assimilação de Dados utilizado pelo Centro Nacional de Previsão (NCEP) e o do European Center for Medium-Range Weather Forecast (ECMWF), que é uma organização intergovernamental composta por 34 países, e pela National Aeronautics and Space Administration (NASA), norte-americana.

1.1.2.1 ECMWF

O ECMWF (European Center for Medium-Range Weather Forecast) coleta informações meteorológicas de estações espalhadas por todo mundo, radares meteorológicos, satélites entre outras fontes. Segundo Moraes et al. (2012), estas diferentes fontes são usadas para estimar os elementos meteorológicos médio diários. Os dados são processados pela empresa alemã Meteo-Consult e posteriormente transferidos para o JRC (Joint Research Centre), que os coloca à disposição de usuários, via internet.

As características atmosféricas são descritas por meio de equações e os dados são processados para formar um modelo fisicamente válido da atmosfera em pontos distantes 0,25 graus de latitude e longitude, resultando em variáveis para o mundo inteiro (ECMWF, 2015; PERSON; GRAZZIANI, 2007).

Blain et al. (2006) compararam dados de temperatura do ar e precipitação obtidos de estações de superfície com dados do modelo ECMWF, no estado de São Paulo, apresentando resultados satisfatórios e que após ajustes, foram usados em um modelo de estimativa de produtividade de soja. O mesmo ocorreu no Rio Grande do Sul quando Melo et al. (2007) aplicaram os dados diretamente sobre um modelo agrometeorológico espectral de produtividade da soja. Para o cultivo do café são escassas as informações que abrangem as principais regiões cafeeiras. Contudo a utilização exclusiva do modelo atmosférico ECMWF, pode gerar alguns erros devido às variabilidades geográficas presentes em diferentes regiões, gerando assim, a falta de acurácia para a região desejada (MORAES et al., 2012).

1.1.2.2 NASA

Outro banco de dados online foi desenvolvido pela NASA, do modelo derivado de dados meteorológicos por satélite em tempo quase real (NASA, 2007; CHANDLER et al., 2010). Os dados são obtidos a partir da Goddard Earth Observation System (GEOS). Estes são derivados de diversas fontes combinadas, como: observações da superfície terrestre, observações de superfície do oceano, radares espaciais, radiossondagens, balões meteorológicos e satélites (WHITE et al., 2008).

A resolução espacial é de 1 ° x 1 ° de latitude e longitude (NASA, 2007; CHANDLER et al., 2010). Alguns trabalhos que utilizaram e avaliaram os dados da NASA são encontrados na literatura, como modelagem de fenologia do trigo nos USA (WHITE et al., 2008), simulação do potencial de produção de milho na China (BAI et al., 2010), previsão de extrema incêndios na Sérvia (WESTBERG et al., 2010), e simulação de índice de satisfação do requisito de água para milho na província de Free State of South Africa (MOELETSI; WALKER, 2012).

1.1.3 Cultivo do Café

O café é um produto de grande importância em todo o mundo exercendo grande influência no mercado internacional (VICTORINO, CARVALHO e FERREIRA, 2016). É cultivado em mais do que 80 países, com uma produção em 2014 de 143 milhões de sacas (ICO, 2015). O Brasil tem a maior produção de café do mundo (AMARASINGHE et al., 2015).

Originário da Etiópia, o cafeeiro pertence à família das Rubiaceas sendo uma planta perene de clima tropical, muito sensível aos elementos meteorológicos, principalmente em escala microclimática (CRAPARO et al., 2015), como radiação solar, temperatura do ar e precipitação pluvial.

De acordo com Camargo e Camargo (2001), a faixa de temperatura ideal para o cultivo do café arábica fica entre 19 e 22 ºC. Temperaturas mais altas promovem formação de botões florais e estimulam o crescimento dos frutos. Entretanto, estimulam também, a proliferação de pragas e aumenta o risco de infecções que podem comprometer a qualidade da bebida. O cafeeiro é pouco tolerante ao frio, sendo que temperaturas abaixo de 10 ºC inibem o crescimento da planta e temperaturas de –2 ºC próxima às folhas já provoca início de danos aos tecidos (DAMATTA, 2007).

Outro agravante é a ocorrência de geadas. Camargo (1975) classificou as geadas quanto aos danos nas lavouras cafeeiras em três categorias: geadas moderadas, geadas severas e geadas severíssimas. As geadas moderadas são aquelas que acontecem uma a cada três anos, elas podem causar grandes prejuízos se não forem observadas as condições microclimáticas, como por exemplo, evitar os plantios nas áreas mais baixas, para onde o ar frio escorre durante as noites. As geadas severas, (uma a cada cinco a seis anos) provocam perdas significativas, caso a lavoura não esteja plantada em condições de topoclima e microclima adequadas. As geadas severíssimas, (três em cem anos) são eventos extremos, que felizmente ocorrem com frequência muito baixa.

Por outro lado, regiões onde ocorrem temperaturas médias anuais superiores a 23°C associadas à seca na época do florescimento provocam abortamento floral e formação de "estrelinhas", o que implica na quebra de produção, principalmente nos anos em que a estação seca mostra-se mais longa ou atrasada (CAMARGO, 1985; THOMAZIELLO et al., 2000).

Quanto às necessidades hídricas do cafeeiro, torna-se difícil estabelecer um padrão ótimo anual de precipitação pluvial, pois depende também de outros fatores, principalmente da distribuição dessas chuvas ao longo do ano. De acordo com Costa (2016) a precipitação ótima está entre 1200 e 1800 mm, mas o cafeeiro cresce e produz bem tanto na precipitação de 800 mm, como em locais onde a precipitação é maior que 2000 mm anuais. Na região do cerrado Mineiro, em razão da ocorrência de período seco durante a estação fria, adota-se na maioria das lavouras, a prática da irrigação (RONCHI et al., 2015; FERNANDES et al., 2016). Esta irrigação, entretanto, deve ser interrompida em alguns casos no período que antecede a florada, pois pesquisas apontam um atraso no desenvolvimento do botão floral, em relação as lavouras não irrigadas (DAMATTA et al., 2007; SILVA et al., 2009; FERNANDES et al., 2016). De acordo com Thomaziello et al. (2000), o cafeeiro arábica tolera bem deficiências hídricas de até 150 mm ano^{-1}, entretanto pesquisas mais recentes, como a de Fernandes et al. (2016) apontam que déficits hídricos superiores a 150 mm, afetam o desenvolvimento vegetativo e produtivo do cultivo.

Tendo em vista que o café arábica está condicionado a alterações durante o seu desenvolvimento, nos seus diversos estádios fenológicos, alterações estas causadas pelas condições meteorológicas, em especial pela disponibilidade hídrica e de temperatura, um estudo de interações clima-produção pode ser realizado, com o uso de modelos agrometeorológicos que procuram quantificar os efeitos das variações meteorológicas na safra (ROSA et al., 2010). Logo, um cálculo para estimar a produção dessa *commodity* torna-se uma ferramenta importante para apoiar o estabelecimento de políticas cafeeiras a fim de orientar os produtores.

1.2 REFERÊNCIAS BIBLIOGRÁFICAS

ALBERGEL, C.; DE ROSNAY, P.; BALSAMO, G.; ISAKSEN, L.; MUÑOZ SABATER, J. Soil moisture analyses at ECMWF: evaluation using global ground-based in-situ observations, **J. Hydrometeorol**., v.13, p.1442–1460, 2012.

AMARASINGHE, U. A.; HOANH, C. T.; D'HAEZE, D.; HUNG, T.Q. Toward sustainable coffee production in Vietnam: More coffee with less water. **Agricultural Systems**, v.136, p.96–105, 2015.

BAI, J.; CHEN, X.; DOBERMMAN, A.; YANG, H.; CASSMAN, K.G.; ZHANG, F. Evaluation of NASA satellite- and model-derived weather data for simulation of maize yield potential in China. **Agronomy Journal**, v. 102, n.1, p. 9–16, 2010.

BLAIN, G. C. et al., Relação entre os dados meteorológicos ECMWF preparados pelo JRC-MARS-FOOD e dados de superfície para estimativa de produtividade de soja no estado de São Paulo, Brasil. International Workshop on Crop Monitoring and Forecasting in South America, 2., 2006. **Proceedings…** Montevideo: South America Scientific Network on Crop Monitoring and Forecasting. 2006. p. 61-68

CAMARGO, A. P. características das geadas excepcionalmente severas de julho de 1975. Congresso Brasileiro de Pesquisas Cafeeiras. **Resumos...**Curitiba, 1975. p.250-252.

CAMARGO, A. P. Florescimento e frutificação de café arábica nas diferentes regiões cafeeiras do Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v.20, n.7, p. 831-839, 1985.

CAMARGO, A.P.; CAMARGO, M.B.P. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**, Campinas, v.60, n.1, p.65-68, 2001.

CAMARGO, M. B. P.; SANTOS, M. A.; PEDRO JUNIOR, M. J.; FAHL, J.I.; BRUNINI, O.; MEIRELES, E. J. L.; BARDIN, L. Modelo agrometeorológico de monitoramento e de estimativa de quebra de produtividade como subsidio à previsão de safra de café (Coffea arabica L.): resultados preliminares. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 2003, Porto Seguro. **Anais...** Brasília, Embrapa Café 7576.

CARVALHO, L. G. de; SEDIYAMA, G.C.; CECON, P.R.; ALVES, H.M.R. Aplicação da análise Harmônica por séries de Fourier para a previsão de produtividade da cultura do café no Estado de Minas Gerais. **Eng. Agrícola**, Jaboticabal, v.25, n.3, p.732-741, set./dez. 2005.

CHANDLER, W. S.; HOELL, J. M.; WESTBERG, D.; WHITLOCK, C. H.; ZHANG, T.; STACKHOUSE JR., P. W. 2010. **Near real-time global radiation and meteorology web services available from NASA**. retrieved on 08.01.2016. Disponível em: <http://www.ases.org/papers/ 120.pdf>. Acesso em 20 ago 2016.

COSTA, J. O. **Padrões de resposta termal ao déficit hídrico na cultura do café irrigado por gotejamento.** 99f. Dissertação (Mestrado). Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2016.

CRAPARO, A. C. W.; VAN ASTEN, P. J. A.; LÄDERACH, P.; JASSOGNE, L. T. P.; GRABA, S. W. Coffea arabica yields decline in Tanzania due to climate change: Global implications. **Agricultural and Forest Meteorology**, v. 207, p. 1-10, 2015.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal Plant Physiology**. Campos dos Goytacazes, v. 18, p. 55-81, 2007.

ECMWF 2015: European Centre for Medium-Range Weather Forecasts, 2015. Disponível em: www.ecmwf.int. Acesso em 5 out. 2016.

FERNANDES, A. L. T.; TAVARES, T. O.; SANTINATO, F.; FERREIRA, R. T.; SANTINATO, R. Viabilidade técnica e econômica da irrigação localizada do cafeeiro, nas condições climáticas do planalto de Araxá, MG. **Coffee Science**, Lavras, v. 11, n. 3, p. 347-358, jul./set. 2016.

GASPAR, N. A.; APARECIDO, L. E. O; ROLIM, G. S.; SOUZA, P. S.; BOTELHOS, T. G. Software for the management of weather stations and for agrometeorological calculations. **Australian Journal of Crop Science**. V. 9, n.6, 2015.

HOLMES, T. R. H. T. J.; JACKSON, R. H.; REICHLE; J. B. BASARA. An assessment of surface soil temperature products from numerical weather prediction models using ground-based measurements. **Water Resour**. V.48, 2012. Disponível em: <10.1029/2011WR010538>. Acesso em 15 set 2016.

IBGE 2016. Instituto Brasileiro de Geografia e Estatística. Disponível em: http://www.sidra.ibge.gov.br/. Acesso em: 06 oct. 2016.

INTERNATIONAL COFFEE ORGANIZATION. **Trade statistics**, 2015. Disponível em: < http://www.ico.org/prices/poproduction> Acesso em: 21 dez. 2015.

JOHANN, J. A.; ROCHA, J. V.; ROLIM, G. S.; LAMPARELLI, R. A. C./ DUFT, D. G. Comparação entre dados de temperatura média de estações meteorológicas com dados do modelo ECMWF entre 2000 e 2009 no Estado do Paraná. XV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO - SBSR, **Anais...** Curitiba, PR, Brasil, 30 de abril a 05 de maio de 2011, INPE p.5817.

KLERING, E.V.; FONTANA, D.C.; BERLATO, M.A.; CARGNELUTTI FILHO, A. Modelagem agrometeorológica do rendimento de arroz irrigado no Rio Grande do Sul. **Pesquisa Agropecuária Brasileira**, v.43, p.549-558, 2008. Disponível em: <10.1590/ S0100-204X2008000500001>

MELO, R. W.; FONTANA. D. C. Estimativa do rendimento de soja usando dados do modelo do ECMWF em um modelo agrometeorológico-espectral no Estado do Rio Grande do Sul. Simpósio Brasileiro de Sensoriamento Remoto, 13., 2007, Florianópolis. **Anais..**. São José dos Campos: INPE, 2007. P. 279-286. Disponível em: http://marte.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2006/11.14.16.16/doc/279-286.pdf Acesso em: 24 set. 2015.

MOELETSI, M. E; WALKER, S. Evaluation of NASA Satellite and Modelled Temperature Data for Simulating Maize Water Requirement Satisfaction Index in the Free State Province of South Africa. **Physics and Chemistry of the Earth,** Parts A/B/C, p.50-52 p.157-164, 2012. Disponível em: http://dx.doi.org/10.1016/j.pce.2012.08.012

MORAES, R.A.; ROCHA, J.V.; ROLIM, G.S.; LAMPARELLI, R.A.; MARTINS, M.S. Avaliação dos dados decendiais de precipitação e temperatura máxima e mínima do ar simulados pelo modelo ECMWF para o estado de São Paulo. **Irriga**, Botucatu, v. 17, n. 3, p. 397-407, 2012.

MORAES, R. A.; ROCHA, J. V.; LAMPARELLI, R. A. C. Determination of total accumulated rainfall, global radiation, evapotranspiration and degree-days originated from the ECMWF model to sugar cane crop. Engenharia Agrícola, Jaboticabal, v.34, n. 2, p. 322-334, 2014.

MORETO, V. B.; ROLIM, G. S. Agrometeorological models for groundnut crop yield forecasting in the Jaboticabal, São Paulo State region, Brazil. **Acta Sci., Agron.** v. 37, n. 4, Maringá Oct./Dec. 2015.

NASA 2007: NASA Surface Meteorology and Solar Energy: Methodology. last viewed ago. 2006. Disponível em: http://www.ceoe.udel.edu/windpower/ResourceMap/SSE_Methodology.pdf. Acesso em 10 set. 2016.

OLIVEIRA, L. J. C.; COSTA, L. C.; SEDIYAMA, G. C.; FERREIRA, W. P. M.; OLIVEIRA, M. J. de. Modelos de estimativa de produtividade potencial para as culturas do feijão e do milho. **Engenharia na agricultura**, Viçosa -MG, v.19 n.4, 2011.

PEREIRA, A.R.; ANGELOCCI, L.R.; SENTELHAS, P.C. 2002: **Agrometeorologia: fundamentos e aplicações práticas.** Guaíba: Agropecuária, 478p.

PERSON, A.; GRAZZIANI, F.: **User guide to ECMWF forecast products**. Meteorological Bulletin M3.2, version 4, 2007.

ROLIM, G. S. de; NOVO, M. C. de S. S.; PANTANO, A. P.; TRANI, P. E. Modelagem agrometeorológica para estimação do desenvolvimento e da produção de jiló. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v.15, n. 8, p. 832-837, 2011.

RONCHI, C. P. et al. Respostas ecofisiológicas de cafeeiros submetidos ao deficit hídrico para concentração da florada no Cerrado de Minas Gerais. **Pesquisa Agropecuária Brasileira**, Brasília, v. 50, n. 1, p. 24-32, jan. 2015.

ROSA, V. G. C.; et al., Estimativa da produtividade de café com base em um modelo agrometeorológico espectral. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 12, p. 1478-1488, 2010.

RUBEL, F; RUDOLF, B. Global daily precipitation estimates proved over the European Alps, **Meteorol. Z.**, v.10, p.403 – 414, 2001.

SANTOS, M. A. dos (2005) **Parametrização de coeficientes de sensibilidade e teste de modelos agrometeorológicos de estimativa de quebra de produtividade do cafeeiro (Coffea arabica L.).** 2006. 146p. Dissertação (Mestrado) - Instituto Agronômico de Campinas, APTA, Campinas, 2005.

SANTOS, M. A.; CAMARGO, M. B. P. Calibration of an agrometeorological model for predicting coffee (*Coffea arabica* L.) productivity in Sao Paulo state, Brazil. **Bragantia**, v.65, n.1, p.173-183, 2006.

SILVA, E. A. et al. Influência de déficits hídricos controlados na uniformização do florescimento e produção do cafeeiro em três diferentes condições edafoclimáticas do Estado de São Paulo. **Bragantia**, Campinas, v. 68, p. 493-501, 2009.

SILVA-FUZZO, D.F. da; PRELA-PANTANO, A.; CAMARGO, M.B.P. de. Modelagem agrometeorológica para estimativa de produtividade de soja para o Vale do Médio Paranapanema-SP. **Irriga**, v.20, p.490-501, 2015. Disponível em: <10.15809/irriga.2015v20n3p490>.

THOMAZIELLO, R. A.; FAZUOLI, L. C.; PEZZOPANE, J. R. M.; FAHL, J. I.; CARELLI, M. L. C. **Café arábica: "Cultura e técnicas de produção"**: Boletim técnico 187. Campinas, Instituto Agronômico, 2000. p.2-6.

VICTORINO, E. C.; CARVALHO, L. G.; FERREIRA, D. F. Modelagem agrometeorológica para a previsão de produtividade de cafeeiros na região sul do estado de Minas Gerais. **Coffee Science**, Lavras, v. 11, n. 2, p. 211-220, abr/jun. 2016.

WANG, X.; SONG, L.; CAO, Y. 2012; Analysis of the Weighted mean temperature of China based on sounding and ECMWF reanalysis data**. Acta Meteorologica Sinica** v.26, n.5, p. 642-652, 2012.

WESTBERG, D.; SOJA, A.; STACKHOUSE JR., P.W. Linking satellite-derived fire counts to satellite-derived weather data in fire prediction models to forecast extreme fires in Siberia. **Geophysical Research Abstracts 12**, EGU2010-5597, 2010.

WHITE, J.F.; HOOGENBOOM, G.; STACKHOUSE JR., P.W.; HOELL, J.M. Evaluation of NASA satellite- and assimilation model-derived long-term daily temperature data over the continental US. **Agricultural and Forest Meteorology** v.148, p.1574– 1584, 2008.

CAPÍTULO 2 - Evaluation of air temperature and rainfall from ECMWF and NASA gridded data for southeastern Brazil

ABSTRACT

The study of climatic variables in large scales with surface meteorological stations is limited due to the low density of these stations in many regions, possible sources of errors related to missing data and uncertainties about the calibration sensors. Global gridded data systems (GD) can minimize these problems. Thus studies that validate GDs with "ground truth" are important for several applications such as climate change. The objective of this study was to compare long series of surface data with decennial estimates of average air temperature (T) and precipitation (P) using data from the European Center for Medium-Range Weather Forecast (ECMWF) and the National Aeronautics and Space Administration (NASA) for important agricultural locations in the states of Minas Gerais and São Paulo in Brazil. The GD performance was evaluated by linear regression analysis. Analyses were performed for each meteorological variable for entire years and separated by seasons. The estimates of T from both ECMWF and NASA systems were accurate with the minimum Willmott concordance index (d) and RMSEp of 0.86, 0.37°C respectively, and precision with R^2 0.61. The estimates of P had a minimum R^2, d and RMSEp, of 0.48, 0.79, 2.15 °C respectively. The decreasing orders of (R^2) were autumn > winter > spring > summer for T and winter > autumn > spring > summer for P, varying from 0.93 to 0.61 for T and from 0.77 to 0.48 for P. The ECMWF system was generally more accurate than the NASA systems for the study area.

KEYWORDS: Climatology, Big data, General circulation model, Ground truth, Tropical climate.

Avaliação dos dados de temperatura do ar e precipitação, provenientes do ECMWF e da NASA para o Sudeste do Brasil

RESUMO

O estudo de variáveis climáticas em grandes escalas com estações meteorológicas de superfície é limitado devido à baixa densidade dessas estações em muitas regiões. Possíveis fontes de erros, estão relacionados a dados faltantes e incertezas sobre os sensores de calibração. Sistemas globais de dados em gride (GD) podem minimizar esses problemas. Assim, estudos que validam os GDs com "ground-thrue" são importantes para várias aplicações, como a mudança climática. O objetivo deste estudo foi comparar séries longas de dados de superfície com estimativas decendiais da temperatura média do ar (T) e de precipitação (P) utilizando dados do Centro Europeu de Previsão Meteorológica de Médio Prazo (ECMWF) e da Administração Nacional de Aeronáutica e do Espaço (NASA) para locais importantes agrícolas nos estados de Minas Gerais e São Paulo no Brasil. O desempenho dos GDs foi avaliado por análise de regressão linear. As análises foram realizadas para cada variável meteorológica durante anos inteiros e separadas por estações. As estimativas de T de ECMWF e NASA foram precisas com o índice mínimo de de Willmott (d) e RMSEp de 0,86, 0,37 ° C, respectivamente, e precisão com R^2 0,61. As estimativas de P tiveram um d mínimo, R^2 e RMSEp, de 0,48, 0,79, 2,15 °C, respectivamente. As ordens decrescentes de (R^2) foram outono> inverno> primavera> verão para T e inverno> outono> primavera> verão para P, variando de 0,93 a 0,61 para T e de 0,77 a 0,48 para P. O sistema ECMWF foi mais preciso que os sistemas da NASA para a área de estudo.

Palavras-chave: Big data, Climatologia, Clima tropical, Modelos de circulação global, verdade do solo.

2.1 INTRODUCTION

Gridded data (GD) combine information for land surfaces and oceans from a variety of sources, including remote sensing. They estimate and forecast the weather, separating the planet into grids (BECHTOLD et al., 2008). Several GDs have been developed, for example by the European Center for Medium-Range Weather Forecasts (ECMWF) and the Global Modeling and Assimilation Office (GMAO) of the National Aeronautics and Space Administration (NASA). The use of meteorological data from GD has recently increased in Brazil, mainly used to crop monitoring (GASPAR et al., 2015). Brazil does not have a meteorological network stations that can fill all country meteorological data needs. Stations are more concentrated on developed areas than remote areas (PEREIRA et al., 2002).

Another limitation of surface data is the high cost and uncertainty due to measurement and connection errors in data loggers, clogging of rain sensors, broken wires and the general lack of sensor calibration (WU et al., 2005; MORAES et al., 2014). According to Moeletsi and Walker (2012), data missing from climatic series is another factor limiting various studies, hindering the development of strategies for the management of natural resources. Another common situation is a poor siting of stations, according to Davey and Pielke (2005), the sensor is in a relatively open location, but the surface under and around the sensor is a patchwork of different land-cover types. Frequently, lawn, asphalt, gravel, bare dirt, and concrete were all in close proximity to each other. The use of meteorological data from GDs such as ECMWF and NASA can resolve these problems.

ECMWF and NASA collect information from various meteorological sources, such as meteorological radar stations around the world and satellites (ECMWF 2015; COUTO et al., 2015; WHITE et al., 2008; NASA 2007; CHANDLER et al., 2010). Data from the ECMWF system are processed and made available in grids of up to 0.25 degrees of latitude and longitude, corresponding to about 27 km in the meridional direction (ECMWF 2015). The NASA system has a higher spatial resolution of 1 × 1°, corresponding to 108 km in the meridional direction, with data collected by the Goddard Earth Observation System (GEOS) (NASA 2007).

Modeled general atmospheric images have been extensively studied by Albergel et al. (2012) in Australia, Holmes et al. (2012) in Oklahoma (USA), Wang et al. (2012) in China, Rubel and Rudolf (2001) in the European Alps and Johann et al. (2011) in Brazil. The use of the ECMWF's ten-day data to identify locations where data is limited or surface data is not available has been proven by some authors as, Verkade et al. (2013), for operational hydrologic forecasting at various spatial scales. Moraes et al. (2014), for determination of total accumulated rainfall, global radiation, evapotranspiration and degree-days to sugar cane crop.

Several studies have used and evaluated data from NASA, such as modeling wheat phenology in the USA (White et al. 2008), simulating corn yield in China (BAI et al., 2010), forecasting fire in Serbia (WESTBERG et al., 2010) and simulating water requirements for corn in the province of Free State in South Africa (MOELETSI; WALKER, 2012). White et al., (2008) and McPhee and Margulis (2005) verified the accuracy of the NASA systems for estimating temperature and precipitation, respectively, both in the USA.

The purpose of our study was to compare and evaluate decennial data of average air temperature and precipitation from the ECMWF and NASA GDs with surface data for southeastern Brazil.

2.2 MATERIAL AND METHODS

2.2.1 Meteorological Stations

Maximum, minimum and average surface air temperatures were collected from September 1992 to December 2015, and rainfall data were collected from June 2000 to January 2016. These data were organized on a scale of 10 days and stratified by season: spring (SP) 22 September to 21 December; summer (SUM) 21 December to 20 March; autumn (AUT) 20 March to 20 June and winter (WIN) 20 June to 22 September.

We used 10 locations (Figure 1). Five of Minas Gerais state: Monte Santo de Minas, Guaxupé, Tiros, Coromandel and Serra do Salitre. Daily information for maximum, minimum and average and air temperatures and rainfall were provided by agricultural cooperatives in the regions. The other five of São Paulo state: Mococa, Gália, Cristais Paulista, Franca and Osvaldo Cruz. Meteorological data were provided by the Agronomic Institute (IAC / APTA), http://www.apta.sp.gov.br/index.php. These locations are not part of large urban centers and are important agricultural areas with coffee, sugarcane, citrus, grains and horticultural crops (IBGE, 2016). Those locations have a total area of 10,014 km^2.

Figure 1. Locations utilized in the study in the state of Minas Gerais and São Paulo, Brazil.

The atmosphere of the region of Cerrado Mineiro (Coromandel, Tiros and Serra do Salitre) is humid based on the climatic classification of Thornthwaite (1948), with moderate water deficits in winter, mesothermal and relative evapotranspiration <48 mm. The other locations in Minas Gerais and São Paulo, except Osvaldo Cruz with a sub-humid climate, have similar climates but no or brief water deficits. The locations at higher altitudes are Serra do Salitre and Franca at 1244 and 1026 m a.s.l., respectively (Table 1).

Table 1. Geographic and climatic description of the locations of the meteorological stations

Regions	Altitude (m)	Climate[1]	Continentality (km)
	Cerrado Mineiro		
Coromandel - MG	928	$B'_2wB'_3a'$	574.24
Tiros - MG	977	$B'_2wB'_3a'$	467.36
Serra do Salitre - MG	1244	$B'_2wB'_3a'$	488.25
	Sul de Minas		
Guaxupé- MG	824	$B'_2rB'_4a'$	288.68
Monte Santo de Minas - MG	892	$B_3rB'_3a'$	316.29
Mococa - SP	665	$B'_2rB'_4a'$	279.19
	Alta Mogiana		
Franca - SP	1026	$B'_3rB'_3a'$	391.01
Cristais Paulista - SP	996	$B'_3rB'_3a'$	413.13
	Noroeste Paulista		
Gália - SP	522	$B'_1rB'_4a'$	343.13
Osvaldo Cruz - SP	440	C_2rAa'	470.75

[1] Following the Thornthwaite climatic classification (1948). [2] Shortest distance to the Atlantic Ocean.

2.2.2 ECMWF

ECMWF is an independent intergovernmental organization that develops, and operates on a regular basis, global meteorological models and data-assimilation systems to forecast the weather based on the input of meteorological data collected by satellites and earth observation systems, such as automatic and manned stations, aircraft, ships and weather balloons. The action of Food Security (FOODSEC) regularly receives outputs daily, for 10-day periods and monthly from the ECMWF global

Operational circulation model (2008-present) and has acquired the archive of the Interim reanalysis model (1989-2012), which has been further processed by Meteoconsult. Data for estimating rainfall, solar radiation, evapotranspiration and temperature were rescaled on a 0.25° grid and aggregated for 10-day periods. More details of the ongoing improvements in resolution can be found at:http://www.ecmwf.int/en/forecasts/documentation-and-support.The maximum, minimum, average and precipitation data are available on a decadal scale for free on the website http://spirits.jrc.ec.europa.eu/.

2.2.3 NASA

Designed primarily for use in architecture, the construction of sustainable buildings and later in agriculture, NASA's POWER project provides daily data for global solar radiation, insolation, surface and air temperatures, relative humidity, wind and precipitation.

Temperature data were from two GMAO models. The first, GEOS-4, provided data from 1 January 1983 to 31 December 2007. The second, GEOS-5, provided data from 1 January 2008 to the present. The types of observations used in GMAO analyses include: observations of Earth's surface and oceans, sea-level winds inferred from backscatter returns from space radars, radiosonde data (e.g. time, temperature, wind and humidity) and air data from meteorological balloons and aircraft and from remotely sensed and satellite information.

The satellite-derived rainfall data set developed by GPCP is available in a 1 × 1° gridded format (GPCP 1DD) from late 1996 to the present. GPCP 1DD is a merged data set based on gauge measurements and satellite estimates of rainfall. Daily data for average temperature and rainfall were downloaded from the NASA/POWER website (STACKHOUSE, 2006): https://power.larc.nasa.gov/.The data were organized in ten-day periods, similar to ECMWF outputs.

2.2.4 Statistical analysis

We analyze the local models (LMs) that were composed by the 10 locations, and the regional models (RMs), those that are the grouping of the location of the same region. Generated by data not seasonally stratified (ALL) and by stratified data, spring

(SPR), summer (SUM), autumn (AUR), winter (WIN). RMs constituted four regions: Sul de Minas (locations: Guaxupé, Monte Santo de Minas and Mococa), Cerrado Mineiro (locations: Coromandel, Tiros and Serra do Salitre), Alta Mogiana (locations: Franca and Cristais Paulista) and Noroeste Paulista (locations Gália and Osvaldo Cruz).

The performances of the data sets derived from ECMWF and NASA were assessed by linear regression analysis, where the independent variables were those from the GDs and the dependent variables were those from the surface stations. A HAT test for identifying outliers and influential points was applied on the observed data. The accuracy were determined using the Willmott et al (1985) concordance index (d) (standardized measure of the degree of model prediction error), and Root Mean Squared Error prediction ($RMSE_P$), (measures degree of non-symmetric error between estimated values and measured/reference data), (MOELESTSI; WALKER, 2012), (Equations 1 and 2). The precision were determined using R^2 (determines the correlation between estimated values and measured/reference data) (FELAYI et al., 2011). (Eq 3).

$$RMSE_P = \sqrt{\sum_{i=1}^{N} \frac{(Yobs_i - Yest_i)^2}{N}} \quad (1)$$

$$d = 1 - \frac{\sum_{i=1}^{N} (Yobs_i - Yest_i)^2}{\sum_{i=1}^{N} (|Yest_i - \bar{Y}| + |Yobs_i - \bar{Y}|)} \quad (2)$$

$$R^2 = \frac{\sum_{i=1}^{N} (Yest_i - \bar{Y})^2}{(Yobs_i - \bar{Y})^2} \quad (3)$$

Where: $Yest_i$ is a variable estimated in year i, $Yobs_i$ is a variable observed in dekal i, $\bar{Y}$ is a variable estimated average and N is the number of datapoints.

The slope of the regression line is obviously important, as it determines the sensitivity of the calibration function; that is, the rate at which the signal changes with concentration. The uncertainty in the slope is expressed as the standard error (or deviation) of the slope, Sb, and is calculated in terms of the standard error of the regression (Equation. 4). The uncertainty in the intercept is also calculated in terms of the standard error of the regression as the standard error (or deviation) of the intercept (Equation 5).

$$S_b = \frac{SE}{\sqrt{var(x)\ x\ (N-1))}} \quad (4)$$

$$S_a = \sqrt{\frac{1}{N} + \frac{\overline{Xobs}^2}{\sum Xobs^2 x \frac{(\sum Xobs^2)}{N}}} \quad (5)$$

Where: SE is a standard error; Yobs is a variable observed and N is the number of datapoints.

Was used a specific Chow test for subsets of data for comparing RMs and LMs. This test compares the sum of squared residuals of the original regression with the sum of the squares of new regressions from the subsamples (Gujarati and Porter 2011). The null hypothesis (H_0) is that the regressions do not differ. F_{CALC} is then calculated (Equation 6):

$$Fcalc = \frac{\frac{(SQE_G - SQE_{IR})}{k}}{\frac{SQE_{IR}}{(n1+n2-2k)}} \quad (6)$$

If F_{CALC} is larger than the critical, H_0 is rejected and the models differ significantly, but if F_{CALC} is smaller than the critical, H_0 is accepted and the models do not differ significantly. We first compared LM and RM generated for each season. The LMs were then compared to their respective RM.

2.3 RESULTS AND DISCUSSION

Average temperature was highest in Osvaldo Cruz, Gália and Mococa at 22.3 °C and lowest in Guaxupé and Monte Santo de Minas at 19 °C (Figure 2A). Winter temperature was lowest in Gália. The temperature in Mococa was expected to be similar to the temperature in Sul de Minas, due to the proximity of the locations, but was not probably because of the low altitude. Precipitation was lower in Gália and Osvaldo Cruz than the other locations (Figure 2B), with an average annual total of 1237 mm compared to 1540 mm.

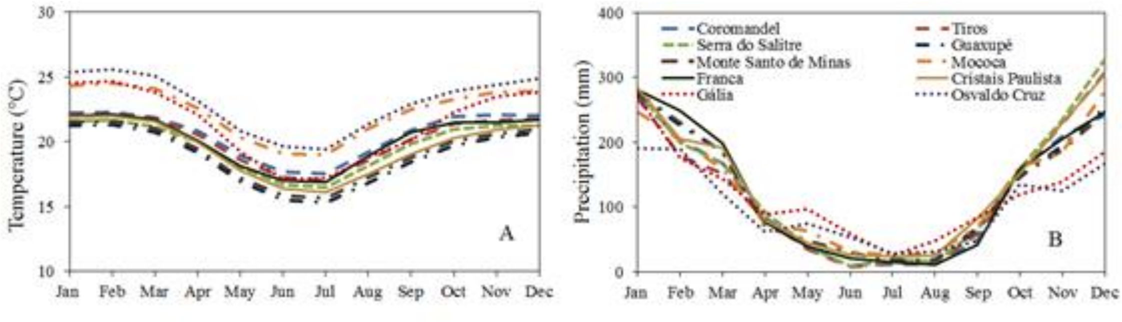

Figure 2. Temperature (T) and precipitation (P) of normal data (INMET, 2016).

All analyzes of the systems (ECMWF and NASA) may present errors associated with the physical inaccuracy of the model numerical methods, background fields, observation data, quality control, and errors generated by data assimilation methods (KUMAR; KISHTAWAL, 2017). This occurs because the systems use different models to produce their first-guess fields and different physical processes to evaluate the climate elements (LIM; HO, 2000). The same authors affirmed that comparisons of different gridded data systems with the 'ground truth' is essential for validation of such systems.

The data for air temperature from the ECMWF systems contained no outliers for any of the locations. The NASA data identified outliers in three locations, one in Guaxupé, two in Tiros and four in Cristais Paulista. The lowest and highest temperatures were 24 and 28 °C, respectively, but the average temperature in these locations was 21 °C.

We identified a larger number of P outliers for both GDs on the observed data, which were removed for all locations. Guaxupé and Osvaldo Cruz both had 13 P outliers. P in Osvaldo Cruz was maximum at 262 mm and averaged 31 mm, and P in Guaxupé was maximum at 255 mm and averaged 38 mm.

The ten-day models using ECMWF data generally overestimated T relative to the surface data; the intercept coefficients were all positive, except for Serra do Salitre (Table 2). The RMSEp values were low for all localities, with a minimum value of 0.37 °C in Serra do Salitre, and a maximum value of 1.97 °C in Cristais Paulista. The values of the standard errors of the coefficients were also low, with average of 0.04 for slope, and 0.92 for intercept. The models differed between seasons in all locations except Franca. The model for Franca was not significant for any season, so the models did not differ from those that used the data for all seasons (ALL).

Spring was significant for all locations except Serra do Salitre, i.e. the SPR model differed significantly from ALL. Gália and Monte Santo de Minas differed only in the autumn. The RMs from the ECMWF data were predominantly significant for the seasons and did not differ only for AUT in the Alta Mogiana region, SPR in the Noroeste Paulista region and SPR and SUM in the Cerrado Mineiro region (Table 3). Alta Mogiana was the RM that obtained the lowest mean value of RMSEp, 1.02°C, however the mean value of all regions was 1.24°C. The mean values of the standard errors of the slope and intercept were, 0.035 and 0.77, respectively.

The models generated from NASA data also overestimated T and P, which generally had positive intercept coefficients (Table 4). This result agrees with that by Moeletsi and Walker (2012), who evaluated temperature data from the GPCP systems of NASA and verified a positive bias for minimum temperature, indicating overestimation compared to surface information, in the province of Free State in South Africa. Albergel et al. (2012) reported that soil-moisture data from the ECMWF systems also tended to be overestimated relative to the observed data in various countries, such as the USA, Australia, France, Spain, Italy, Germany, Poland and Finland.

In contrast, Che et al. (2016) found that a systematic negative bias for maximum and minimum temperatures was consistent throughout the maize growing season in northeastern China. These authors, though, also found an increase in the bias between November and March. More research is needed to understand the reasons for these differences. Many surface meteorological stations, however, are near urban areas, where the temperature is normally higher than the surrounding rural areas due to the effect of heat islands (ZHOU et al., 2004). For example, Ren et al. (2007) found evidence of urbanization-induced warming in two meteorological stations in China.

The RMSEp values, for the NASA data were low for all localities, with a minimum value of 0.50°C in Guaxupé, and a maximum value of 2.81°C in Cristais Paulista. The values of the standard errors of the coefficients were also low, with average of 0.06 for slope, and 1.54 for intercept. These results were similar to those found for Moeletsi and Walker (2012), which obtained mean values lower than 4°C.

Most LM estimates of temperature from NASA data differed significantly; only Tiros, Serra do Salitre, Monte Santo de Minas and Gália did not at a maximum of two stations for ALL. The RMs over the seasons also mostly differed significantly, except WIN for Sul de Minas and SPR and WIN for Cerrado Mineiro (Table 5). The minimum RMSEp for RMs with NASA data was also obtained in the Alta Mogina region, with the value of 1.09°C, and the average of all regions was of 1.40°C. The mean values of the standard errors of the slope and intercept were, 0.032 and 0.81, respectively.

Table 2. Models for estimating surface temperature for the seasons from gridded data in the ECMWF systems. The differences between the models for the locations for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	RMSEP	Slope	Sb	Intercept	Sa	Seasons	RMSEP	Slope	Sb	Intercept	Sa
Coromandel MG						Franca SP					
ALL	0.68	0.92	0.02	2.39	0.50	ALL	0.73	0.87	0.02	3.72	0.39
SPR *	0.65	0.79	0.05	5.24	1.23	SPR NS	0.70	0.88	0.05	3.41	1.13
SUM NS	0.65	0.90	0.08	2.90	1.93	SUM NS	0.43	0.83	0.04	4.63	0.94
AUT *	0.63	1.10	0.06	-1.00	1.31	AUT NS	0.56	0.92	0.04	2.65	0.70
WIN NS	0.61	0.87	0.05	3.07	1.07	WIN NS	1.04	0.81	0.06	4.86	1.13
Tiros MG						Cristais Paulista SP					
ALL	0.61	0.92	0.01	2.32	0.30	ALL	0.97	0.97	0.02	1.70	0.38
SPR *	0.70	0.81	0.04	4.95	1.00	SPR *	1.97	0.68	0.07	8.17	1.51
SUM *	0.50	0.64	0.04	8.91	0.95	SUM *	0.63	0.65	0.10	9.07	2.41
AUT *	0.57	1.07	0.04	-0.75	0.73	AUT NS	0.99	0.99	0.04	1.39	0.83
WIN NS	0.52	0.93	0.03	2.13	0.62	WIN *	1.00	0.95	0.04	1.60	0.78
Serra do Salitre MG						Mococa SP					
ALL	0.54	0.88	0.02	1.64	0.41	ALL	0.73	0.97	0.01	-0.33	0.24
SPR *	0.45	1.02	0.04	-1.61	0.90	SPR *	0.67	0.90	0.02	0.74	0.49
SUM NS	0.37	0.98	0.05	-0.71	1.07	SUM *	0.85	0.91	0.03	1.14	0.72
AUT NS	0.44	0.99	0.05	-0.48	1.01	AUT NS	0.61	0.97	0.03	-0.64	0.75
WIN *	0.63	1.00	0.07	-0.47	1.34	WIN *	0.65	0.89	0.02	1.20	0.52
Guaxupé MG						Osvaldo Cruz SP					
ALL	0.62	1.00	0.01	0.95	0.30	ALL	1.11	1.02	0.02	2.28	0.43
SPR NS	0.59	0.89	0.04	3.32	0.96	SPR *	0.96	1.24	0.07	-3.53	1.61
SUM *	0.38	0.72	0.04	7.51	0.85	SUM NS	0.83	1.21	0.09	-2.09	2.11
AUT *	0.49	1.11	0.03	-1.07	0.58	AUT *	0.92	1.19	0.03	-1.01	0.73
WIN *	0.68	0.84	0.05	3.63	0.83	WIN NS	1.22	0.95	0.05	3.76	0.98
Monte Santo de Minas MG						Gália SP					
ALL	0.90	0.76	0.02	6.09	0.46	ALL	0.95	1.17	0.02	-3.38	0.35
SPR *	0.70	1.02	0.05	0.12	1.16	SPR *	0.96	0.99	0.05	0.58	1.22
SUM *	0.47	1.14	0.05	-2.57	1.13	SUM *	0.75	0.93	0.06	2.63	1.40
AUT NS	0.86	0.84	0.05	4.49	0.97	AUT NS	0.75	1.21	0.03	-4.24	0.67
WIN *	0.90	0.86	0.07	4.45	1.33	WIN *	0.90	0.97	0.04	0.10	0.71

***** The models differed significantly at *P*<0.05. **NS** The models did not differ significantly at *P*<0.05.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models.

Table 3. Regional models (RMs) for estimating surface temperature for the seasons from gridded data in the ECMWF systems. The differences between RMs for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa	Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa
Sul de Minas						Alta Mogiana					
ALL	1.20	0.77	0.01	4.74	0.26	ALL	0.84	0.93	0.01	2.37	0.28
SPR *	1.20	1.15	0.03	-3.56	0.67	SPR *	0.82	0.76	0.04	6.16	0.95
SUM *	1.15	0.76	0.04	5.31	0.81	SUM *	1.65	0.69	0.06	8.11	1.30
AUT *	1.02	0.66	0.02	7.50	0.42	AUT NS	0.82	0.96	0.03	1.89	0.58
WIN *	1.18	0.63	0.02	7.67	0.42	WIN *	0.98	0.92	0.03	2.36	0.63
Cerrado Mineiro						Noroeste Paulista					
ALL	0.75	0.93	0.01	1.84	0.31	ALL	1.59	1.21	0.02	-3.17	0.41
SPR NS	0.96	0.89	0.04	2.85	0.95	SPR NS	1.95	1.40	0.05	-8.15	1.15
SUM NS	0.91	0.79	0.06	5.14	1.26	SUM *	1.61	1.40	0.07	-7.60	1.73
AUT *	2.01	1.11	0.04	-1.81	0.78	AUT *	1.65	1.35	0.04	-5.75	0.90
WIN *	0.72	0.97	0.03	1.15	0.64	WIN *	1.88	1.19	0.05	-2.65	1.00

***** The models differed significantly at $P<0.05$. **NS** The models did not differ significantly at $P<0.05$.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models.

Table 4. Models for estimating surface temperature for the seasons from data in the NASA systems. The differences between the models of the locations for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	RMSE$_P$	Slope	Sb	Intercept	Sa	Seasons	RMSE$_P$	Slope	Sb	Intercept	Sa
Coromandel MG						Franca SP					
ALL	0.93	0.99	0.04	3.66	2.56	ALL	1.05	0.67	0.02	6.85	0.51
SPR *	0.86	1.06	0.10	1.06	8.01	SPR *	1.02	0.60	0.06	8.11	1.50
SUM *	0.70	0.83	0.11	5.59	8.93	SUM *	0.57	0.94	0.07	1.30	1.61
AUT *	0.60	0.99	0.11	3.43	2.53	AUT *	0.70	0.83	0.04	3.63	0.86
WIN *	0.81	0.92	0.11	2.35	2.29	WIN *	1.02	0.61	0.04	7.58	0.92
Tiros MG						Cristais Paulista SP					
ALL	0.85	0.62	0.02	8.52	0.42	ALL	1.39	0.65	0.02	7.17	0.48
SPR NS	0.85	0.55	0.10	10.19	2.48	SPR *	2.74	0.24	0.06	17.25	1.37
SUM *	0.77	0.24	0.05	17.90	1.10	SUM *	2.19	0.29	0.08	16.75	1.91
AUT NS	0.66	0.65	0.05	8.04	1.08	AUT *	2.81	0.04	0.04	4.49	0.77
WIN *	0.62	0.62	0.03	8.24	0.51	WIN *	1.05	0.64	0.03	6.03	0.73
Serra do Salitre MG						Mococa SP					
ALL	0.94	0.68	0.03	6.24	0.59	ALL	1.19	0.77	0.01	4.28	0.34
SPR *	0.92	0.83	0.07	2.11	1.74	SPR *	1.03	0.72	0.03	4.79	0.66
SUM *	0.81	1.39	0.18	-8.89	3.98	SUM *	1.09	0.67	0.03	6.26	0.77
AUT NS	0.54	0.79	0.05	4.17	0.96	AUT *	1.07	0.61	0.04	8.62	1.06
WIN NS	0.89	0.64	0.07	6.88	1.32	WIN *	1.18	0.71	0.04	5.81	0.87
Guaxupé MG						Osvaldo Cruz SP					
ALL	0.81	0.91	0.02	2.57	0.38	ALL	1.72	0.78	0.03	6.83	0.63
SPR *	0.73	0.85	0.06	3.34	1.40	SPR *	1.57	0.39	0.06	16.67	1.59
SUM *	0.50	0.91	0.07	3.08	1.56	SUM *	1.17	0.62	0.10	12.13	2.48
AUT *	0.64	1.05	0.04	0.10	0.75	AUT *	1.33	1.02	0.04	2.05	1.01
WIN *	0.69	0.71	0.04	5.87	0.73	WIN *	1.48	0.69	0.05	7.89	1.05
Monte Santo de Minas MG						Gália SP					
ALL	1.07	0.64	0.03	8.70	0.67	ALL	1.48	0.93	0.02	0.87	0.48
SPR *	0.84	0.98	0.07	0.37	1.52	SPR *	1.44	0.65	0.07	7.52	1.40
SUM *	1.51	1.13	0.12	-2.17	2.78	SUM *	1.09	0.58	0.07	10.48	1.68
AUT NS	1.06	0.66	0.05	8.26	1.00	AUT NS	1.12	0.98	0.04	-0.05	0.79
WIN NS	1.11	0.64	0.09	8.57	1.57	WIN *	1.25	0.72	0.04	0.72	0.83

***** The models differed significantly at P<0.05. **NS** The models did not differ significantly at P<0.05. ***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models.

Table 5. Regional models (RMs) for estimating surface temperature for the seasons from data in the NASA systems. The differences between RMs for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa	Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa
Sul de Minas						Alta Mogiana					
ALL	1.33	0.68	0.01	6.72	0.27	ALL	1.24	0.65	0.02	7.08	0.37
SPR *	1.12	0.95	0.02	0.57	0.50	SPR *	1.49	0.39	0.06	13.86	1.40
SUM *	1.49	0.50	0.04	11.15	0.89	SUM *	0.71	0.66	0.05	8.20	1.20
AUT *	1.11	0.62	0.02	8.49	0.48	AUT *	0.86	0.88	0.04	2.54	0.73
WIN NS	1.34	0.60	0.02	8.49	0.48	WIN *	1.17	0.65	0.04	6.17	0.79
Cerrado Mineiro						Noroeste Paulista					
ALL	0.94	0.67	0.01	7.25	0.30	ALL	2.00	0.97	0.02	1.08	0.47
SPR NS	3.48	0.68	0.04	6.52	0.94	SPR *	2.03	0.70	0.06	7.59	1.97
SUM *	0.99	0.56	0.05	10.26	1.09	SUM *	1.65	0.84	0.07	5.31	1.76
AUT *	0.75	0.77	0.03	5.24	0.60	AUT *	1.81	1.14	0.04	-2.11	0.89
WIN NS	0.77	0.63	0.03	7.71	0.63	WIN *	1.85	0.90	0.04	1.81	0.80

***** The models differed significantly at $P<0.05$. **NS** The models did not differ significantly at $P<0.05$.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models.

The models were generally accurate; minimum d was 0.53 for NASA data in Tiros for SUM. The good correlation was consistent with those in other areas that clearly show the utility of global meteorological data from NASA satellite derivatives (WHITE et al., 2008; BAI et al., 2010; MOELETSI; WALKER, 2012). The model generated using ECMWF data, however, was better, with d above >0.8, whereas d for the NASA systems was >0.6 (Figure 3A).

The NASA system was also less precise, especially in the spring, summer and autumn, in five locations: Gália, Osvaldo Cruz, Cristais Paulista, Monte Santo de Minas and Tiros. The weather and topography in the gridded areas of the NASA figures are possible sources of errors (CHE et al., 2016). White et al. (2008) reported higher differences in air temperature from measurements at NASA ground stations in mountainous and coastal areas.

Other locations had an R^2 >0.5 (Figure 3B). Moeletsi and Walker (2012) reported similar results, with R^2 between 0.69 and 0.79. Che et al. (2016), though, reported an average R^2 of 0.9 for NASA and surface data. The ECMWF systems had a minimum precision of 0.6 and an average of 0.81.

Both models were more accurate and precise for all seasons. The ECMWF and NASA systems had average values of 0.96 and 0.91 and an average R^2 of 0.87 and 0.74, respectively. These results indicate that the temperature data for both models could be used when the temperature data from the meteorological stations are limited or in areas without meteorological stations.

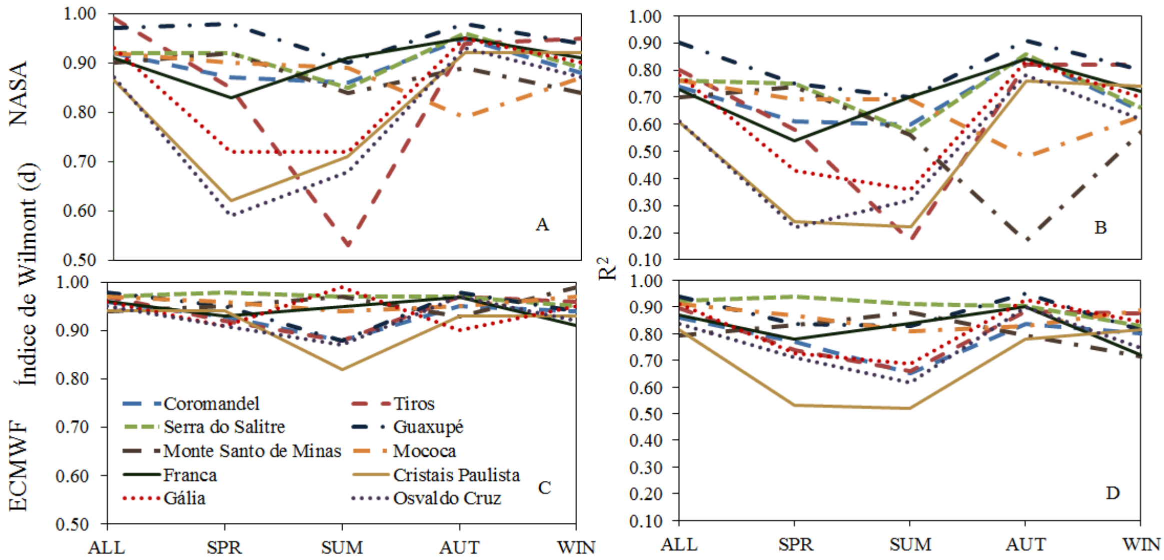

Figure 3. Accuracy (d) and precision (R^2) of temperature estimated by local models (LM) using NASA (A and B) and ECMWF (C and D) data.

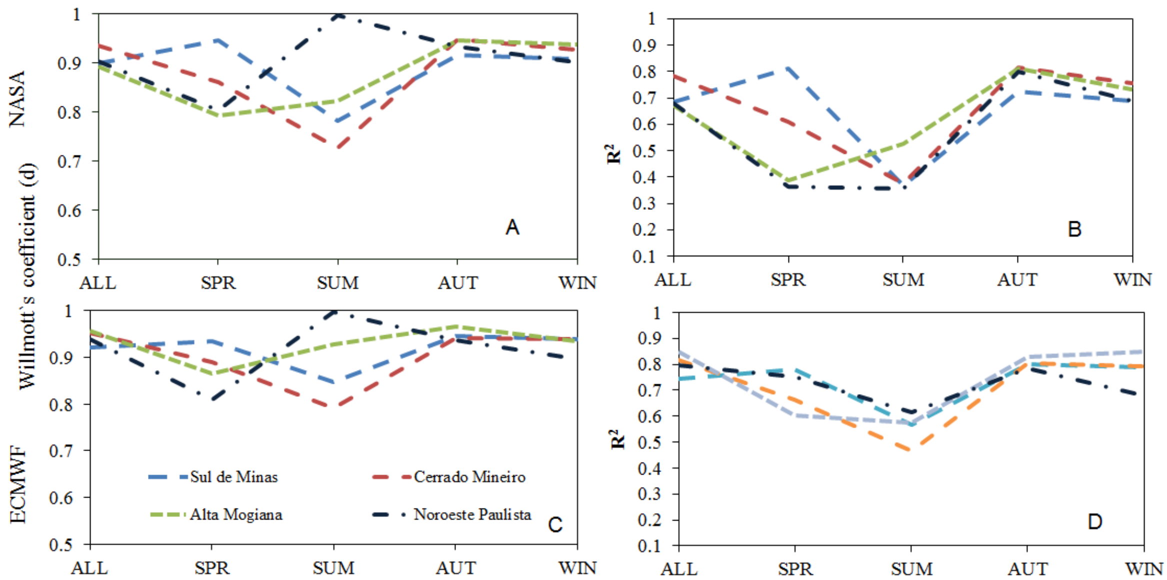

Figure 4. Accuracy (d) and precision (R^2) of temperature estimated by regional models (RMs) using NASA (A and B) and ECMWF (C and D) data.

The RM and LM for all data models (ALL) estimating temperature using ECMWF and NASA data were also overestimated relative to the surface data, except for the ECMWF models for the Noroeste Paulista region, which had a mean intercept coefficient of -2.9474 (Table 6). Only models for the Alta Mogiana region did not differ significantly, because other LMs differed from their respective region.

Table 6. Regional and local models of temperature using the ECMWF and NASA data. The differences between the regional models and locations were analyzed by a Chow test. Legend: SM – Sul de Minas; GXP – Guaxupé; MSM – Monte Santo de Minas; MOC – Mococa; AM – Alta Mogiana; FRA – Franca; CPA – Cristais Paulista; CM – Cerrado Mineiro; COR – Coromandel; TIR – Tiros; SSA – Serra do Salitre; NP – Noroeste Paulista; GAL – Gália; OSC – Osvaldo Cruz;RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa	Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa
ECMWF											
SM	1.20	0.77	0.01	4.74	0.26	**CM**	0.75	0.93	0.01	1.84	0.31
GXP *	0.62	1.00	0.01	0.95	0.30	COR*	0.68	0.92	0.02	2.39	0.50
MSM *	0.90	0.76	0.02	6.09	0.46	TIR *	0.61	0.92	0.01	2.32	0.30
MOC *	0.73	0.97	0.01	-0.33	0.24	SSA *	0.54	0.88	0.02	1.64	0.41
AM	0.75	0.93	0.01	1.84	0.31	**NP**	1.59	1.21	0.02	-3.17	0.41
FRA NS	0.73	0.87	0.02	3.72	0.39	GAL *	0.95	1.17	0.02	-3.38	0.35
CPA NS	0.97	0.97	0.02	1.70	0.38	OSC*	1.11	1.02	0.02	2.28	0.43
NASA											
SM	1.33	0.68	0.01	6.72	0.27	**CM**	0.94	0.67	0.01	7.25	0.30
GXP *	0.81	0.91	0.02	2.57	0.38	COR*	0.93	0.99	0.04	3.66	2.56
MSM *	1.07	0.64	0.03	8.70	0.67	TIR *	0.85	0.62	0.02	8.52	0.42
MOC *	1.19	0.77	0.01	4.28	0.34	SSA *	0.94	0.68	0.03	6.24	0.59
AM	1.24	0.65	0.02	7.08	0.37	**NP**	2.00	0.97	0.02	1.08	0.47
FRA NS	1.05	0.67	0.02	6.85	0.51	GAL *	1.48	0.93	0.02	0.87	0.48
CPA NS	1.39	0.65	0.02	7.17	0.48	OSC*	1.72	0.78	0.03	6.83	0.63

***** The models differed significantly at $P<0.05$. **NS** The models did not differ significantly at $P<0.05$.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models

The ECMWF global circulation systems using ten-day period data predominantly overestimated rainfall relative to the surface data. This result agrees with Lim and Ho 2000, who compared rainfall data observed with the assimilation products of ECMWF, NCEP/NCAR, and NASA-GEOS-, they verified that the ECMWF overestimates the precipitation over tropical oceans as compared to the surface data. The intercept coefficients were positive, except for AUT in Tiros and WIN in Serra do

Salitre (Table 7). Rainfall was also overestimated using the NASA data, and only WIN at Serra do Salitre and Cristais Paulista were underestimated (Table 9). Rahman and Sengupta (2007) reported that daily precipitation data from three satellites (TRMM 3B42-V5 and 3B42-V6, GPCP and IMD) overestimate the media observed.

There was a higher variation in the RMSEp values for the precipitation, and the seasons, SUM and SPR, presented superior values, whereas AUT and WIN, inferior values. Kishore et al. (2016) found a similar result, these authors validated the precipitation data from India with observations and reanalysis datasets and spatial trends. They found that, the JRA-25 data reanalysis heavily underestimates during the pre-monsoon season. These differences are only found in spatial distribution, but monthly mean values show reasonably good agreement.

The minimum value of the RMSEp was 2.15 mm in WIN of Coromandel, and the maximum was 46.63 mm in SUM of Osvaldo Cruz. However, the mean value of RMSEp was 27.16 mm. The mean values of the standard errors of the slope and intercept were, 0.09 and 4.10, respectively. Noroeste Paulista was the RM that obtained the lowest mean value of RMSE, 26.48 °C, however the mean value of all regions was 27.31 °C. The mean values of the standard errors of the slope and intercept were, 0.055 and 2.63, respectively.

Only three stations did not differ from ALL models for their respective locations to the ECMWF data: SPR in Tiros, WIN in Osvaldo Cruz and AUT in Gália. The others were significant, so the models differed. The same was true for the RMs where SUM in Sul de Minas, SPR in Cerrado Mineiro and AUT and WIN in Noroeste Paulista differed significantly from ALL models (Table 8). Those from NASA models behaved similarly, where the majority was not significant, except for SUM in Coromandel and Gália, SPR in Franca, WIN in Gália and AUT in Serra do Salitre, Mococa, Osvaldo Cruz and Gália. The opposite occurred for RM; most models were significant, except for AUT in Sul de Minas, SPR in Alta Mogiana and SUM in Cerrado Mineiro (Table 10).

The minimum value of the RMSEp, for the NASA data, was 5.53 mm in AUT of Serra do Salitre, and the maximum was 40.24 mm in AUT of Mococa. However, the mean value of RMSEp was 26.17 mm, this value is inferior than that obtained with the data from ECMWF. The mean values of the standard errors of the slope and intercept were, 0.09 and 3.97, respectively. Noroeste Paulista was the RM that obtained the lowest mean value of RMSE, 25.09 °C, however the mean value of all regions was 27.33 °C. The mean values of the standard errors of the slope and intercept were, 0.042 and 2.53, respectively.

Table 7. Models for estimating surface rainfall for the seasons using gridded data in the ECMWF systems. The differences between the models of the locations for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa	Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa
Coromandel MG						Franca SP					
ALL	33.35	1.06	0.05	4.64	1.05	ALL	26.79	0.84	0.04	6.93	2.23
SPR ^{NS}	38.44	1.03	0.11	10.49	6.66	SPR ^{NS}	43.71	0.92	0.13	11.67	8.71
SUM ^{NS}	55.85	1.11	0.19	10.03	13.79	SUM ^{NS}	38.32	0.80	0.10	18.69	8.16
AUT ^{NS}	20.37	0.62	0.10	10.15	4.08	AUT ^{NS}	19.48	0.83	0.09	5.72	3.08
WIN ^{NS}	2.15	1.06	0.10	0.35	0.28	WIN ^{NS}	10.02	0.95	0.13	2.18	1.30
Tiros MG						Cristais Paulista SP					
ALL	27.39	0.95	0.03	3.70	1.61	ALL	26.42	0.90	0.04	4.52	1.86
SPR *	27.35	0.71	0.06	9.78	3.63	SPR ^{NS}	36.57	0.75	0.09	17.40	6.35
SUM ^{NS}	42.49	0.97	0.08	10.29	6.27	SUM ^{NS}	40.50	0.75	0.10	18.35	8.47
AUT ^{NS}	26.04	1.09	0.08	-0.32	3.13	AUT ^{NS}	15.13	0.73	0.07	2.58	2.09
WIN ^{NS}	4.43	0.66	0.06	0.41	0.42	WIN ^{NS}	11.06	1.33	0.13	0.48	1.34
Serra do Salitre MG						Mococa SP					
ALL	28.39	1.27	0.05	4.55	2.70	ALL	27.49	0.86	0.03	7.28	1.63
SPR ^{NS}	31.52	1.42	0.14	4.86	7.22	SPR ^{NS}	9.01	0.86	0.06	0.98	0.93
SUM ^{NS}	44.84	1.22	0.13	13.25	10.61	SUM ^{NS}	19.94	0.71	0.07	8.71	2.26
AUT ^{NS}	5.63	0.94	0.09	6.98	3.67	AUT ^{NS}	22.47	0.94	0.10	5.64	2.71
WIN ^{NS}	3.26	1.75	0.07	-0.17	0.52	WIN ^{NS}	27.00	0.89	0.07	4.76	3.76
Guaxupé MG						Osvaldo Cruz SP					
ALL	24.30	0.93	0.04	4.01	2.00	ALL	25.15	0.69	0.04	8.33	1.81
SPR ^{NS}	28.60	0.83	0.09	10.74	5.45	SPR ^{NS}	29.46	0.82	0.10	6.46	5.28
SUM ^{NS}	36.14	0.70	0.10	21.38	8.76	SUM ^{NS}	45.63	0.74	0.12	12.10	10.11
AUT ^{NS}	23.04	1.02	0.09	1.32	4.02	AUT ^{NS}	21.85	0.75	0.12	5.02	3.22
WIN ^{NS}	8.41	0.49	0.07	2.93	1.09	WIN *	13.25	1.53	0.12	3.44	1.62
Monte Santo de Minas MG						Gália SP					
ALL	30.58	1.10	0.04	4.36	2.39	ALL	98.52	0.86	0.03	5.34	1.58
SPR ^{NS}	33.78	0.97	0.11	10.38	6.71	SPR ^{NS}	27.41	0.78	0.08	10.71	4.28
SUM ^{NS}	42.81	1.03	0.11	15.55	10.08	SUM ^{NS}	40.63	0.82	0.10	11.87	7.68
AUT ^{NS}	27.17	0.95	0.10	4.71	4.31	AUT *	16.12	0.61	0.07	6.36	2.04
WIN ^{NS}	8.92	0.52	0.10	2.45	1.18	WIN ^{NS}	10.60	0.89	0.06	2.40	1.08

***** The models differed significantly at *P*<0.05. **NS** The models did not differ significantly at *P*<0.05.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models

Table 8. Regional models (RMs) of surface precipitation for the seasons using data in the ECMWF systems. The differences between RMs for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	RMSEP	Slope	Sb	Intercept	Sa	Seasons	RMSEP	Slope	Sb	Intercept	Sa
Sul de Minas						Alta Mogiana					
ALL	28.02	0.93	0.02	5.95	1.14	ALL	28.26	0.87	0.03	6.56	1.51
SPR NS	23.03	0.97	0.04	4.18	1.81	SPR NS	38.53	0.75	0.08	17.70	5.40
SUM *	34.83	0.91	0.05	10.71	3.04	SUM NS	38.46	0.76	0.07	18.71	5.98
AUT NS	32.62	0.87	0.05	9.15	3.11	AUT NS	17.19	0.78	0.06	3.53	1.82
WIN NS	19.57	0.88	0.04	3.08	1.44	WIN NS	11.16	1.05	0.11	1.95	1.03
Cerrado Mineiro						Noroeste Paulista					
ALL	28.88	1.00	0.02	4.67	1.22	ALL	28.62	0.83	0.02	6.57	1.33
SPR *	34.63	0.84	0.06	11.96	3.27	SPR NS	28.24	0.79	0.06	9.00	3.31
SUM NS	50.25	1.06	0.07	11.89	5.69	SUM NS	42.57	0.78	0.08	12.12	6.14
AUT NS	24.88	1.01	0.05	2.45	2.17	AUT *	20.74	0.67	0.07	6.30	1.95
WIN NS	3.53	1.21	0.04	-0.06	0.24	WIN *	12.21	1.11	0.07	3.28	0.96

***** The models differed significantly at $P<0.05$. **NS** The models did not differ significantly at $P<0.05$.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models

Table 9. Models for estimating surface rainfall to the seasons using data in the NASA systems. The differences between the models of the locations for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	RMSE$_P$	Slope	Sb	Intercept	Sa	Seasons	RMSE$_P$	Slope	Sb	Intercept	Sa
Coromandel MG						Franca SP					
ALL	32.18	0.99	0.04	3.66	2.56	ALL	26.19	0.79	0.03	6.27	2.14
SPR ^{NS}	41.40	1.07	0.99	7.64	8.01	SPR *	32.61	0.94	0.09	5.70	6.12
SUM *	40.42	0.83	0.11	5.59	8.93	SUM ^{NS}	36.46	0.67	0.08	18.84	7.74
AUT ^{NS}	17.08	0.99	0.11	3.43	2.53	AUT ^{NS}	23.11	0.71	0.10	8.49	3.57
WIN ^{NS}	16.45	0.92	0.11	2.35	2.29	WIN ^{NS}	10.59	0.90	0.07	1.26	1.32
Tiros MG						Cristais Paulista SP					
ALL	26.45	0.83	0.03	4.30	1.56	ALL	27.08	0.76	0.03	5.18	1.92
SPR ^{NS}	28.76	0.68	0.06	10.04	3.91	SPR ^{NS}	32.30	0.74	0.08	10.96	5.70
SUM ^{NS}	35.68	0.80	0.06	13.13	5.91	SUM ^{NS}	37.92	0.68	0.07	15.36	7.58
AUT ^{NS}	23.88	0.79	0.07	5.63	2.81	AUT ^{NS}	20.09	0.67	0.07	4.39	2.68
WIN ^{NS}	6.48	0.64	0.07	0.02	0.64	WIN ^{NS}	11.50	0.97	0.08	-0.30	1.40
Serra do Salitre MG						Mococa SP					
ALL	32.69	1.19	0.06	3.77	3.14	ALL	26.25	0.83	0.03	5.82	1.63
SPR ^{NS}	34.88	0.94	0.12	23.04	7.34	SPR ^{NS}	13.11	0.68	0.05	1.81	1.29
SUM ^{NS}	45.06	1.37	0.13	-3.17	11.15	SUM ^{NS}	21.12	0.71	0.06	8.08	2.31
AUT *	5.53	0.83	0.10	4.16	4.14	AUT *	40.24	0.78	0.07	19.52	6.83
WIN ^{NS}	8.87	1.21	0.10	-0.73	1.46	WIN ^{NS}	29.25	0.71	0.06	9.74	3.75
Guaxupé MG						Osvaldo Cruz SP					
ALL	23.97	0.77	0.03	0.77	1.90	ALL	25.56	0.81	0.03	2.64	1.66
SPR ^{NS}	25.78	0.80	0.08	13.37	4.68	SPR ^{NS}	25.78	0.89	0.07	1.31	4.33
SUM ^{NS}	32.22	0.59	0.08	26.08	7.12	SUM ^{NS}	32.71	0.80	0.07	3.27	6.12
AUT ^{NS}	35.03	0.71	0.08	6.27	3.53	AUT *	16.97	0.49	0.07	6.68	2.44
WIN ^{NS}	9.69	0.85	0.05	1.12	1.23	WIN ^{NS}	13.97	0.99	0.07	1.29	1.75
Monte Santo de Minas MG						Gália SP					
ALL	31.28	0.91	0.04	7.14	2.38	ALL	27.96	0.70	0.03	6.46	1.80
SPR ^{NS}	33.69	0.75	0.09	16.34	6.13	SPR ^{NS}	28.47	0.48	0.07	18.99	4.23
SUM ^{NS}	43.33	0.85	0.09	19.52	9.65	SUM *	39.69	0.71	0.08	16.10	6.95
AUT ^{NS}	25.59	0.97	0.09	4.52	3.98	AUT *	21.28	0.45	0.07	8.62	2.67
WIN ^{NS}	14.42	0.70	0.08	2.96	1.80	WIN *	17.33	0.66	0.05	1.69	1.71

***** The models differed significantly at *P*<0.05. **NS** The models did not differ significantly at *P*<0.05.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models

Table 10. Regional models (RMs) of surface precipitation to the seasons from the NASA systems data. The differences between RMs for all data (ALL) and stations were analyzed by a Chow test. Legend: RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	$RMSE_P$	Slope	Sb	Intercept	SEa	Seasons	$RMSE_P$	Slope	Sb	Intercept	Sa
Sul de Minas						Alta Mogiana					
ALL	29.30	0.84	0.02	6.99	1.15	ALL	26.74	0.75	0.02	6.30	1.41
SPR *	24.22	0.81	0.04	6.34	1.83	SPR NS	35.49	0.81	0.07	11.67	4.77
SUM *	33.39	0.83	0.04	9.95	2.87	SUM *	37.84	0.65	0.05	18.28	5.65
AUT NS	33.32	0.87	0.04	9.11	3.01	AUT *	19.44	0.67	0.05	4.68	2.01
WIN *	20.90	0.77	0.03	3.78	1.51	WIN *	10.30	0.87	0.05	0.41	0.92
Cerrado Mineiro						Noroeste Paulista					
ALL	30.49	0.92	0.02	4.24	1.29	ALL	26.62	0.74	0.02	5.19	1.24
SPR *	35.12	0.84	0.05	9.83	3.29	SPR *	30.21	0.64	0.06	12.91	3.32
SUM NS	52.69	0.98	0.07	9.59	6.08	SUM *	39.29	0.70	0.06	15.98	4.89
AUT *	28.53	0.91	0.06	4.36	2.48	AUT *	17.95	0.52	0.05	6.20	1.73
WIN *	3.48	0.65	0.04	0.07	0.24	WIN *	11.37	0.70	0.04	1.30	0.93

***** The models differed significantly at $P<0.05$. **NS** The models did not differ significantly at $P<0.05$.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models

The accuracy of both models was generally satisfactory; minimum d was 0.59 for AUT in Gália using the NASA data. The ECMWF model performed better than the NASA systems, with d >0.8 and >0.6, respectively (Figure 4A and C).These differences may be attributed partially to the different resolutions of the observations (KISHORE et al., 2016). The GPCP (NASA) product does not use information from rain collectors, only satellite data (Huffman et al. 2001). The data varied among the seasons but was more attenuated in the ECMWF systems. The variability increased in autumn and winter in both models. This variability of the data may be due to the topography, changes in land cover or localized effects and may lead to large errors in the assimilation model (WHITE et al., 2008).

The precision of the models was similar, with R^2 ranging from 0.3 to 0.8. R^2 for Serra do Salitre in the ECMWF model was 0.9 for WIN, indicating that the ECMWF systems performed better than the NASA systems (Figure 4B and D). The variability of the data was also similar. Our study supports the results by Rubel and Rudolf (2001) for determining how well global precipitation fields, often used for verifying climatic models, represent reality. A comparison of the performance of the same systems that we used in our study, ECMWF and GPCP – 1DD (NASA), found that both global data sets had deficits for realistically estimating rainfall in India. The ECMWF predictions, however, forecasted the spatial distribution of wet areas better than the NASA predictions. A major finding of the study by Rubel and Rudolf (2001) was that the accuracy of the GPCP-1DD estimates was significantly lower than those of the ECMWF forecasts.

We also noted that the accuracy and precision of the NASA model were lower in Gália but were similar among the locations in the ECMWF model. The accuracy and precision of the precipitation estimates were highest in Serra do Salitre for both models. This location was at the highest elevation (1244 m a.s.l., Table 1), indicating that the ECMWF systems may be best in tropical areas at higher altitudes.

The inability to evaluate the accuracy of the surface data was a limitation of this study. Analyses of limited sets of stations have indicated that station-to-station errors can be as large as the GD errors (WU et al., 2005; MAHMOOD et al., 2006; WHITE et al., 2008). Both data sources have inherent errors and uncertainties. Uncertainties associated with satellite parameters include pixel size, sensor resolution, and navigation time, accuracy of the algorithm and geographic coincidence of information captured by a satellite with information from ground measurements (CHE et al., 2016). For example, Pinker and Laszlo (1992) showed that the differences between ground-based and satellite-derived values can be reduced if the satellite data had a higher resolution, which could also apply to the GDs.

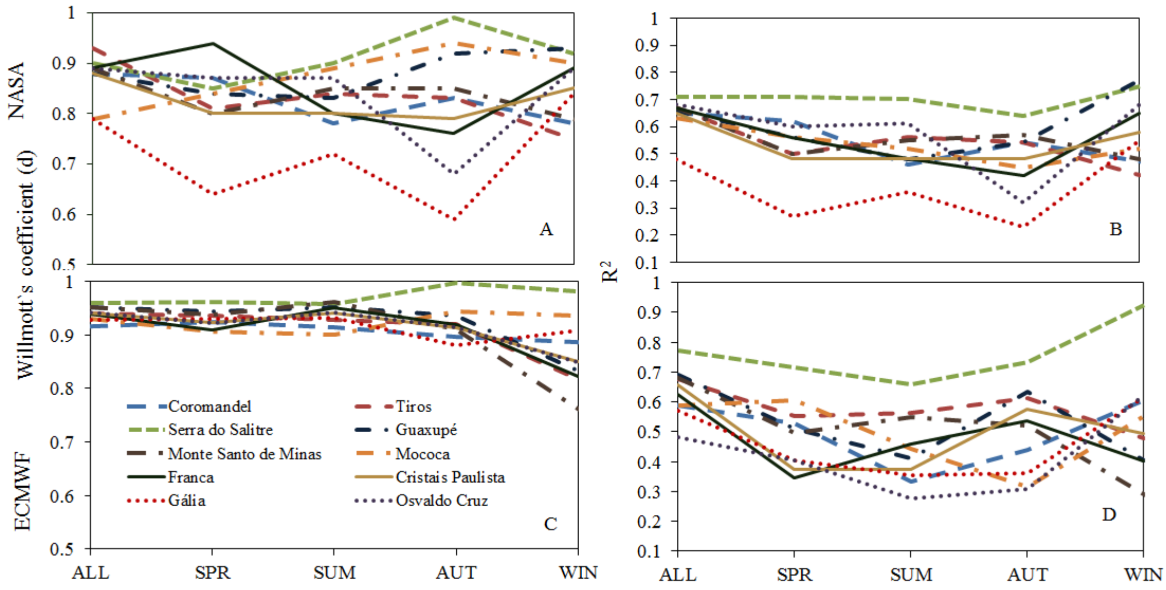

Figure 5. Accuracy (d) and precision (R^2) of the local precipitation models using NASA (A and B) and ECMWF (C and D) data.

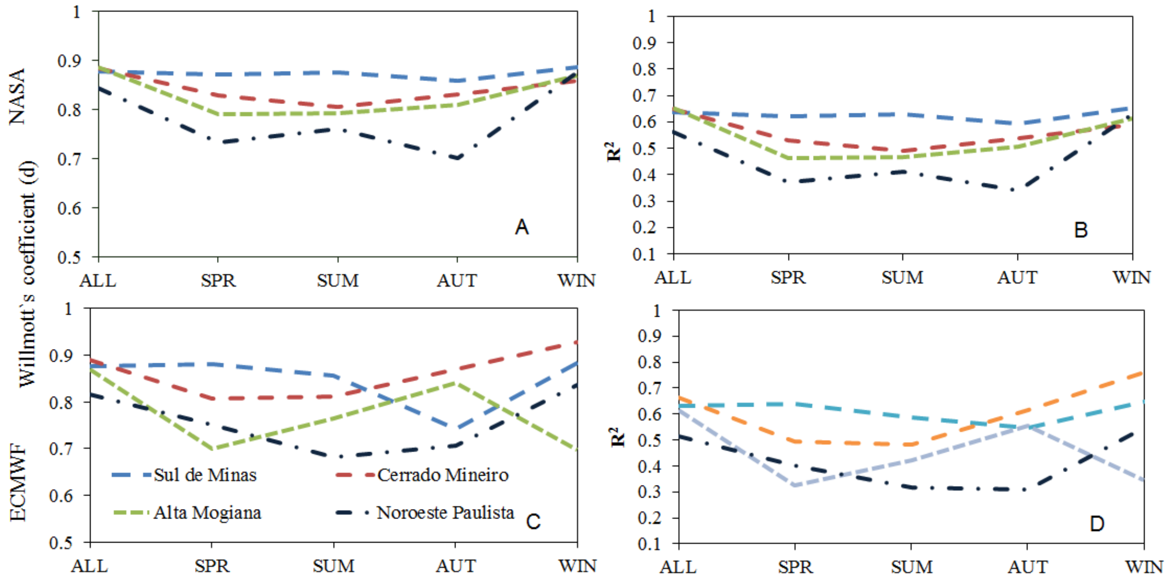

Figure 6. Accuracy (d) and precision (R^2) of the regional precipitation models using NASA (A and B) and ECMWF (C and D) data.

All models, both local and regional, overestimated rainfall relative to the surface data for both data sources, ECMWF and NASA. Only the models for Monte Santo de Minas, Tiros and Serra do Salitre differed from their RMs for the ECMWF data, and the models for Tiros and Serra do Salitre differed from the RM for the NASA data. The models for the other locations did not differ significantly from their RMs (Table 11).

Table 11. Regional models (RMs) and temperature estimates using the ECMWF and NASA data. The differences between RMs and locations were analyzed by a Chow test. Legend: SM – Sul de Minas; GXP – Guaxupé; MSM – Monte Santo de Minas; MOC – Mococa; AM – Alta Mogiana; FRA – Franca; CPA – Cristais Paulista; CM – Cerrado Mineiro; COR – Coromandel; TIR – Tiros; SSA – Serra do Salitre; NP – Noroeste Paulista; GAL – Gália; OSC – Osvaldo Cruz; RMSEp, Root Mean Squared Error prediction; Sb, standard error of the slope; Sa, standard error of the intercept.

Models											
Seasons	RMSE$_P$	Slope	Sb	Intercept	Sa	Seasons	RMSE$_P$	Slope	Sb	Intercept	Sa
ECMWF											
SM	28.02	0.93	0.02	5.95	1.14	**CM**	28.88	1.00	0.02	4.67	1.22
GXP ^{NS}	24.30	0.93	0.04	4.01	2.00	COR^{NS}	33.35	1.06	0.05	4.64	1.05
MSM *	30.58	1.10	0.04	4.36	2.39	TIR *	27.39	0.95	0.03	3.70	1.61
MOC ^{NS}	27.49	0.86	0.03	7.28	1.63	SSA *	28.39	1.27	0.05	4.55	2.70
AM	28.26	0.87	0.03	6.56	1.51	**NP**	28.62	0.83	0.02	6.57	1.33
FRA ^{NS}	26.79	0.84	0.04	6.93	2.23	GAL ^{NS}	98.52	0.86	0.03	5.34	1.58
CPA ^{NS}	26.42	0.90	0.04	4.52	1.86	OSC^{NS}	25.15	0.69	0.04	8.33	1.81
NASA											
SM	29.30	0.84	0.02	6.99	1.15	**CM**	30.49	0.92	0.02	4.24	1.29
GXP ^{NS}	23.97	0.77	0.03	0.77	1.90	COR^{NS}	32.18	0.99	0.04	3.66	2.56
MSM ^{NS}	31.28	0.91	0.04	7.14	2.38	TIR *	26.45	0.83	0.03	4.30	1.56
MOC ^{NS}	26.25	0.83	0.03	5.82	1.63	SSA *	32.69	1.19	0.06	3.77	3.14
AM	26.74	0.75	0.02	6.30	1.41	**NP**	26.62	0.74	0.02	5.19	1.24
FRA ^{NS}	26.19	0.79	0.03	6.27	2.14	GAL ^{NS}	27.96	0.70	0.03	6.46	1.80
CPA ^{NS}	27.08	0.76	0.03	5.18	1.92	OSC^{NS}	25.56	0.81	0.03	2.64	1.66

***** The models differed significantly at P<0.05. **NS** The models did not differ significantly at P<0.05.

***** The test was significant at 5% probability; there is a statistical difference between the models. **NS** The test was not significant at 5% probability; there is no statistical difference between the models

Both models generally best represented the air temperature of the precipitation, because the dispersion of the data was added to the precipitation. We expected this result, because precipitation is very difficult to estimate due to its high spatial variability. Estimating continuous global daily precipitation fields is challenging (RUBEL; RUDOLF, 2001). One of the most challenging problems in climate science is observing and understanding the spatial and temporal variations of tropical rainfall. Rainfall varies on time scales ranging from minutes and hours (diurnal) to weeks, seasons, years and longer. Lorenz and Kunstmann (2012) evaluating the hydrological cycle in three state-of-the-art reanalyses, observed that, the highest differences in spatial variability and the amount of rainfall can be found over tropical South America, and these differences cause large-scale deviation patterns. According to Kishore et al. (2016), the representation of precipitation in the gridded datasets comprises a number of challenging aspects, which may not occur when considering continuous variables such as temperature or pressure, which vary smoothly over a large spatial scale. Furthermore, rain is associated with cloud systems that have complex structures and are organized on different spatial scales. Measuring atmospheric variables over land is even difficult because of their high variability (RAHMAN; SENGUPTA, 2007).

There was no tendency for temperature models in relation to altitude; however, near the altitude of 800m R^2 values were higher for both ECMWF and NASA. For the precipitation, there was a linear tendency, whereas higher to altitude the values of R^2 (Figure 7).

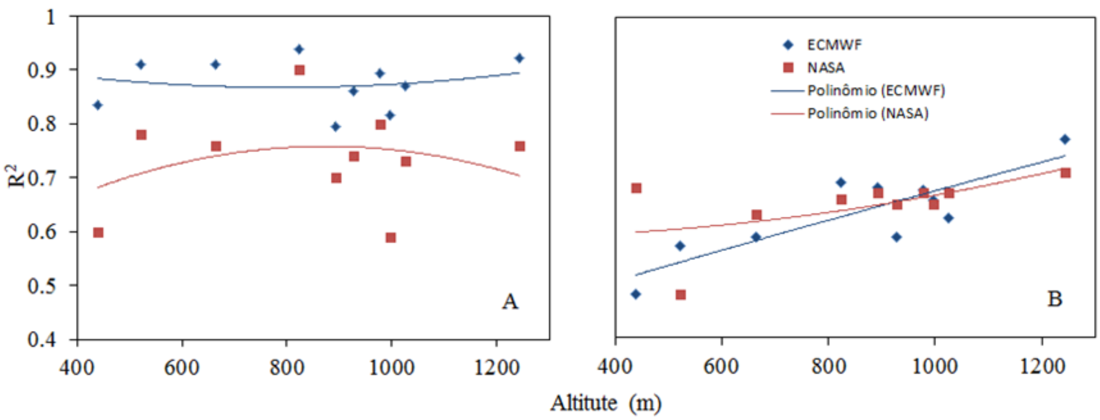

Figure 7. Precision (R^2) of, A) air temperature, and B) precipitation models according to altitude.

The relationship between the models (Figure 8) indicated that the NASA systems overestimated temperatures more than the ECMWF systems, especially in Franca, Cristais Paulista, Gália and Osvaldo Cruz. The dispersion of the data stabilized with the seasons, except for the Sul de Minas region, where spring and autumn were reversed. The opposite occurred for the precipitation data (Figure 9); the NASA systems underestimated precipitation more than the ECMWF systems. The ECMWF systems had fewer scattered temperature datapoints was smaller, perhaps because the available rainfall data were only from 2000.

The data were standardized with the season in all regions. The data were more variable for SUM for all regions, which illustrate one of the limitations of GDs for estimating precipitation, because summer is the rainy season in the Southern Hemisphere. Rubel and Rudolf (2001) also found deficits in precipitation estimation in India. The difficulty with this estimation is due to the variability in precipitation on time scales ranging from minutes and hours (days) to weeks, seasons, years and longer (RAHMAN; SENGUPTA, 2007). Clouds are also associated with systems with complex structures, arranged on different spatial scales. Atmospheric variables are even difficult to measure on land due to their high variability.

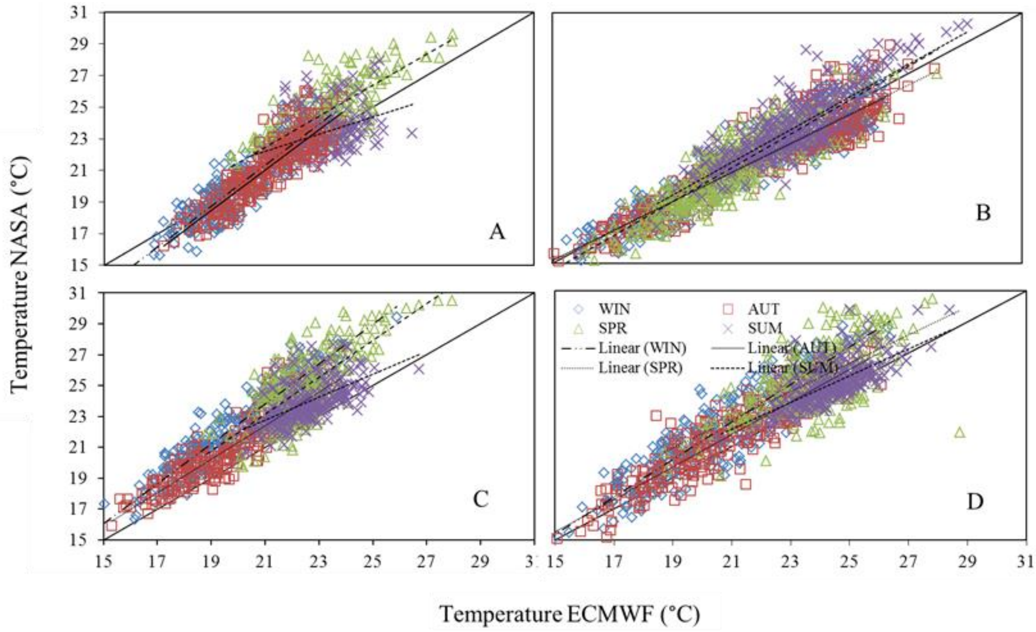

Figure 8. Relationship between temperature (T) estimated by the NASA and ECMWF models for the Cerrado Mineiro and Sul de Minas (A and B) and Alta Mogiana and Noroeste Paulista (C and D) regions.

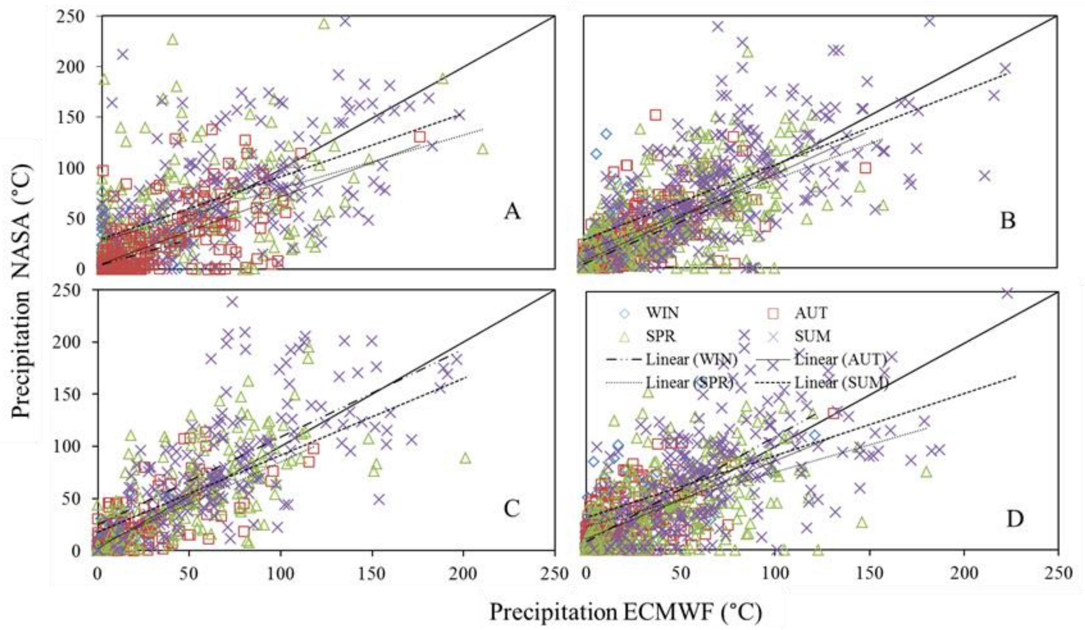

Figure 9. Relationship between precipitation (P) estimated by the NASA and ECMWF models for the Cerrado Mineiro and Sul de Minas (A and B) and Alta Mogiana and Noroeste Paulista (C and D) regions.

2.4 CONCLUSION

The estimates of ten-day periods of air temperatures (T) by both the ECMWF and NASA systems had a minimum R^2 and d for all analysis equal to 0.61 and 0.86, respectively. The T estimated by most models differed from those estimated by the general model, which used all the data (ALL) and regional data. Meaning that one should use the specific models to estimate T for each locations and season. For rainfall (P) the minimum R^2 and d for both systems were 0.48, 0.79, respectively. The most models did not differ from locations and seasons, therefore we can use regional models and/or all season's models.

2.5. REFERENCES

ALBERGEL, C.; DE ROSNAY, P.; BALSAMO, G.; ISAKSEN, L.; MUÑOZ SABATER, J. Soil moisture analyses at ECMWF: evaluation using global ground-based in-situ observations, **J. Hydrometeorol**., v.13, p.1442–1460, 2012.

BAI, J.; CHEN, X.; DOBERMMAN, A.; YANG, H.; CASSMAN, K.G.; ZHANG, F. Evaluation of NASA satellite- and model-derived weather data for simulation of maize yield potential in China. **Agronomy Journal** v.102, n.1, p. 9–16, 2010.

BECHTOLD, P.; KOHLER, M.; JUNG, T.; DOBLAS-REYES, F.; RODWELL, M.J.; VIART, F.; BALSAMO, G. Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decendal time-scales. **Q.J.R. Meteorol. Soc.** v. 134, p. 1337-1351, 2008.

CHANDLER, W. S.; HOELL, J. M.; WESTBERG, D.; WHITLOCK, C. H.; ZHANG, T.; STACKHOUSE JR., P. W. 2010. **Near real-time global radiation and meteorology web services available from NASA**. retrieved on 08.01.2016. Disponível em: <http://www.ases.org/papers/ 120.pdf>. Acesso em 20 ago 2016.

CHE, T.; DAI, L.; ZHENG, X.; LI, X.; ZHAO, K. Estimation of snow depth from passive microwave brightness temperature data in forest regions of northeast China. **Remote Sensing of Environment**. V.183, p.334-349, 2016.

COUTO, F. T.; SALGADO, R.; COSTA, M. J.; PRIOR, V. Precipitation in the Madeira Island over a 10-year period and the meridional water vapour transport during the winter seasons. **Int. J. Climatol.** V.35, p.3748–3759, 2015.

DAVEY, C. A.; PIELKE, R. A. S. R. Microclimate exposures of surface-based weather stations–implications for the assessment of long-term temperature trends. **Bulletin of the American Meteorological Society,** n.86, p. 497–504, 2005.

ECMWF 2015: European Centre for Medium-Range Weather Forecasts, 2015. Disponível em: www.ecmwf.int. Acesso em 5 out. 2016.

GASPAR, N. A.; APARECIDO, L. E. O; ROLIM, G. S.; SOUZA, P. S.; BOTELHOS, T. G. 2015: Software for the management of weather stations and for agrometeorological calculations. **Australian Journal of Crop Science** v. 9, n.6, 2015.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**, 5. ed. Bookman: Porto Alegre. HOLDER, C., BOYLES, R., SYED, A., NIYOGI, D., RAMAN, S., 2006: Comparison of collocated automated (NCECONet) and manual (COOP) climate observations in North Carolina. J. Atmos. Ocean. Technol. V.23, p.671–682, 2011.

HOLDER, C.; BOYLES, R.; SYED, A.; NIYOGI, D.; RAMAN, S. Comparison of collocated automated (NCECONet) and manual (COOP) climate observations in North Carolina. **J. Atmos. Ocean. Technol.,** v. 23, p.671–682, 2006.

HOLMES, T. R. H. T. J.; JACKSON, R. H.; REICHLE; J. B. BASARA. An assessment of surface soil temperature products from numerical weather prediction models using ground-based measurements. **Water Resour**. V.48, 2012. Disponível em: <10.1029/2011WR010538>. Acesso em 15 set 2016.

HUFFMAN, G.J.; MORRISSEY, M.; BOLVIN, D.T.; CURTIS, S.; JOYCE, R.; MCGAVOCK, B.; SUSSKIND, J. Global precipitation at one degree daily resolution from multisatellite observations. **Journal of Hydrometeorology**, v.2, p. 36–50, 2001.

IBGE 2016. Instituto Brasileiro de Geografia e Estatística. Disponível em: http://www.sidra.ibge.gov.br/. Acesso em: 06 oct. 2016.

INMET (2016): Instituto Nacional de Meteorologia. Normais climatológicas do Brasil 1961-1990. Disponível em: http://www.inmet.gov.br/portal/index.php?r=clima/normaisClimatologicas, Acess in 3 abr. 2016.

JOHANN, J. A.; ROCHA, J. V.; ROLIM, G. S.; LAMPARELLI, R. A. C./ DUFT, D. G. Comparação entre dados de temperatura média de estações meteorológicas com dados do modelo ECMWF entre 2000 e 2009 no Estado do Paraná. XV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO - SBSR, **Anais...** Curitiba, PR, Brasil, 30 de abril a 05 de maio de 2011, INPE p.5817

KISHORE, P.; JYOTHI, S.; BASHA, G.; RAO, S. V. B.; RAJEEVAN, M.; VELICOGNA I. Precipitation climatology over India: validation with observations and reanalysis datasets and spatial trends. **Clim. Dyn.,** v. 46, n.1, p. 541–556, 2016, doi: 10.1007/s00382-015-2597-y.

KUMAR, P.; KISHTAWAL, C. M.; PAL, P. K. Impact of ECMWF, NCEP, and NCMRWF global model analysis on the WRF model forecast over Indian Region. **Theoretical Appl Climatology,** v. 127, p.143-151, 2017, doi: 10.1007/s00704-015-1629-1.

LIM, H. S.; HO, C. H. Comparison of tropical rainfall between the observed GPCP data and the assimilation products of ECMWF, NCEP/NCAR, and NASA-GEOS-1. **J. Meteor. Soc. Japan**, v. 78, p. 661–672, 2000.

LORENZ, C.; KUNSTMANN, H. The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis.**J Hydrometeorol,** v.13, p. 1397–1420, 2012.

MAHMOOD, R.; FOSTER, S.A.; LOGAN, D. The GeoProfile metadata, exposure of instruments, and measurement bias in climatic record revisited. **Int. J. Climatol**. V.26, p.1091–1124, 2006.

MCPHEE, J.; MARGULIS, S. A., 2005: Validation and Error Characterization of the GPCP-1DD Precipitation Product over the Contiguous United States. **J. Hydrometeor,** v. 6**,** p.441–459.

MOELETSI, M. E; WALKER, S. Evaluation of NASA Satellite and Modelled Temperature Data for Simulating Maize Water Requirement Satisfaction Index in the Free State Province of South Africa. **Physics and Chemistry of the Earth,** Parts A/B/C, p.50-52 p.157-164, 2012. Disponível em: http://dx.doi.org/10.1016/j.pce.2012.08.012

MORAES, R. A.; ROCHA, J. V.; LAMPARELLI, R. A. C. Determination of total accumulated rainfall, global radiation, evapotranspiration and degree-days originated from the ECMWF model to sugar cane crop. **Eng. Agríc., Jaboticabal** , v. 34, n. 2, p. 322-331, 2014.

NASA 2007: NASA Surface Meteorology and Solar Energy: Methodology. last viewed ago. 2006. Disponível em: http://www.ceoe.udel.edu/windpower/ResourceMap/SSE_Methodology.pdf. Acesso em 10 set. 2016.

PEREIRA, A.R.; ANGELOCCI, L.R.; SENTELHAS, P.C. 2002: **Agrometeorologia: fundamentos e aplicações práticas**. Guaíba: Agropecuária, 478p.

PINKER, R.T.; LASZLO I. Modeling surface solar irradiance for satellite applications on a global scale. **J. Appl. Meteorol**. v31, p.194–211, 1992.

RAHMAN, H.; SENGUPTA, D. Preliminary comparison of daily rainfall from satellites and Indian gauge data. **CAOS Technical** report n. 2007 AS1.

REN, G.Y.; CHU, Z.Y.; CHEN, Z.H.; REN, Y.Y. Implications of temporal change in urban heat island intensity observed at Beijing and Wuhan stations. **Geophys. Res.** V.34, p.L05711, 2007.

RUBEL, F; RUDOLF, B. Global daily precipitation estimates proved over the European Alps, **Meteorol. Z.,** v.10, p.403 – 414, 2001.

STACKHOUSE, P. Prediction Of Worldwide Energy Resources. http://power.larc.nasa.gov (verified March 24,2008). (2006)

THORNTHWAITE, C. W. Na approach toward a rational classification of climate. **Geographical Review**, v. 38, p. 55-94, 1948.

VERKADE, J. S.; BROWN, J.; REGGIANI, P.; WEERTS, A. H. Post-processing ECMWF precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at various spatial scales. **J. Hydrol.,** v. 501, p. 73–91, 2013, doi:10.1016/j.jhydrol.2013.07.039

WANG, X.; SONG, L.; CAO, Y. Analysis of the Weighted mean temperature of China based on sounding and ECMWF reanalysis data. **Acta Meteorologica Sinica** v.26, n.5, p.642-652, 2012.

WESTBERG, D., SOJA, A., STACKHOUSE JR., P.W. Linking satellite-derived fire counts to satellite-derived weather data in fire prediction models to forecast extreme fires in Siberia. **Geophysical Research Abstracts**, v.12, EGU2010-5597, 2010.

WHITE, J.F., HOOGENBOOM, G., STACKHOUSE JR., P.W., HOELL, J.M. Evaluation of NASA satellite- and assimilation model-derived long-term daily temperature data over the continental US. **Agricultural and Forest Meteorology** v.148, p.1574– 1584, 2008.

WILLMOTT, C. J. On the validation of models, **Physical geography**, v. 2, p.184-194, 1981.

WU, H.; HUBBARD, K.G.; YOU, J. Some concerns when using data from the Cooperative Weather Station Networks: a Nebraska case study. **J. Atmos. Ocean. Technol.** V.22, p.592–602, 2005.

ZHOU, L.M.; DICKINSON, R. E.; TIAN, Y.H.; FANG, J.Y.; LI, Q.X.; KAUFMANN, R.K.; TUCKER, C.J.; MYNENI, R.B. Evidence for a significant urbanization effect on climate in China. **Proc. Natl. Acad. Sci**. USA, v.101, p.9540–9544, 2004.

CAPÍTULO 3 - Estimation of C*offea Arabica* yield from Gridded Data

ABSTRACT

Agrometeorological models for estimating yield have been applied to monitoring crops, providing information for internal and external market planning. A limiting factor of these models is the origin of the input data, which are mostly from surface meteorological stations (SMS). We propose the use of meteorological data from gridded data (GD) as a viable and innovative alternative. GDs use data grids of combinations of various sources, including weather balloons, remote sensing, and even SMS. The great advantages over using data from SMS are the easy access to data, the ability to use well-known systematic errors and the extensive coverage of the entire planet. The objective of this study was to estimate coffee yield in important coffee-producing regions of the Brazilian states of São Paulo and Minas Gerais using a modified agrometeorological model initially developed by Santos and Camargo (2006) (Y_{SC}) with GD provided and developed by the European Center for Medium-Range Weather Forecasts (ECMWF) and the National Aeronautics and Space Administration (NASA). The original Y_{SC} model was based on penalties for adverse temperatures (maximum and minimum), water deficiency and yield in the previous year. We proposed modifications to the model to the penalties for temperature extremes and the coefficients calibrated from the different data sources. Mean absolute percentage error (MAPE), systematic root mean square error (RMSEs) and non-systematic root mean square error (RMSEn) decreased relative to the models that used SMS ($Y_{SC\text{-}S}$), indicating that the calibration of the coefficients for the ECMWF and NASA data was efficient. Mean MAPEs were 23.76 and 24.61% for the calibrations with ECMWF and NASA data, respectively. The average tendency was approximately 5 sacks^{-1} for ECMWF and 6 sack^{-1} for NASA. The mean accuracies were 3.11 and 4.39 sacks ha^{-1} for ECMWF and NASA, respectively. The tests with independent data confirmed the improvement of the calibration; MAPE, RMSEs and RMSEn were lower than for the models calibrated with SMS. These results indicate that the models were accurate, precise and had a low tendency.

Key words: Air temperature, Coffee, Crop model, General circulation model, Risk analysis, Water deficiency.

Estimativa da produtividade de *Coffea arabica* a partir de Dados em Grid

RESUMO

A utilização de modelos agrometeorológicos de estimativa de produtividade tem ganhado destaque no monitoramento das safras, proporcionando o planejamento dos mercados interno e externo. Um fator limitante desses modelos é a origem dos dados de entrada, sendo que em sua maioria são provenientes de estações meteorológicas de superfície (SMS). Avaliamos neste trabalho a variação da acurácia do modelo de Santos e Camargo (2006) (Y_{SC}) para estimação de produtividade de *Coffea arabica* a partir de dados em grid (GD) para o sudeste brasileiro. Os GDs fornecem dados em grids resultantes de combinações de diversas fontes, entre eles balões meteorológicos, sensoriamento remoto e até mesmo de SMSs. A grande vantagem em relação às SMSs é o fácil acesso e a obtenção de um erro sistemático mais conhecido, além da ampla cobertura espacial abrangendo todo o planeta. Este trabalho teve como objetivo estimar a produtividade de café pelo modelo agrometeorológico de Santos e Camargo (2006) (Y_{SC}), utilizando dados meteorológicos em grid fornecidos e desenvolvidos pelo European Centre for Medium-Range Weather Forecasts (ECMWF) e pela National Aeronautics and Space Administration (NASA) e por SMS, em regiões cafeeiras importantes dos estados de São Paulo e Minas Gerais. O Y_{SC} baseia-se na penalização por temperaturas adversas (máximas e mínimas), deficiência hídrica e produtividade do ano anterior. Propusemos novas parametrizações no modelo em relação às penalizações por extremos de temperatura e aos coeficientes a serem calibrados decorrentes das diferentes fontes de dados. A calibração dos coeficientes para dados do ECMWF e NASA foi eficiente, pois houve diminuição do MAPE, RMSEs e RMSEu em relação aos modelos que usaram dados de SMS ($Y_{SC\text{-}S}$). Para a calibração com dados do ECMWF, o MAPE médio foi igual a 23,76%. O modelo da NASA obteve uma média de 24,61% de erro. A tendência média (RMSEs) foi de aproximadamente 5 sc ha^{-1} para ECMWF e 6 sc ha^{-1} para a NASA. A precisão média (RMSEn) foi de 3,11 sc ha^{-1} e 4,39 sc ha^{-1} para ECMWF e NASA, respectivamente. Os testes com dados independentes confirmaram a melhoria da calibração, pois houve uma redução do MAPE, RMSEs e RMSEn em relação aos modelos calibrados com dados de SMS.

Palavras - Chave: Café, crop model, general circulation models, risk analysis, big data, deficiência hídrica, temperatura do ar.

3.1 INTRODUCTION

Coffee is the second most valuable *commodity* on the world market, second only to oil (DAMATTA, 2007). It is the most consumed drink in the world and contains several functional compounds, such as caffeine, amino acids and phenols (BUTT; SULTAN, 2011; KITZBERGER et al., 2013). Knowledge of agricultural crops, by estimating regional yields, has become a strategic issue for countries for the planning of internal supplies and international markets (SANTOS; CAMARGO, 2006).

Climatic conditions have the most influence on crop production (WAGNER et al., 2011). Coffee plants (*Coffea arabica*) are very sensitive to meteorological elements, mainly at the microclimatic scale (CRAPARO et al., 2015), such as solar radiation, air temperature and precipitation (HOOGENBOOOM, 2000). Agrometeorological models are one way to measure the effects of climate on crop yield (ROLIM et al., 2008).

Tosello and Arruda (1962), Camargo et al. (1984), Silva et al. (1987), Liu and Liu (1988), Weill (1990), Piccini et al. (1999), Carvalho et al. (2003), Camargo et al. (2003) and Santos and Camargo (2006) have used agrometeorological models to estimate coffee yield in Brazil. Modeling has also been used in other important areas of the world, such as by Dauzat (2001) in Costa Rica and Montoya et al. (2009) and Rodriguez et al. (2011) in Colombia. Modeling studies basically estimate yield from the meteorological conditions during the productive cycle, taking into account the climatic needs of the crop. The model developed by Santos and Camargo (2006) is most commonly used in Brazil to estimate coffee yield. This model is based on the model developed by Camargo et al. (2003) that used agrometeorological data to estimate losses of yield in coffee production based on phenological, hydric and thermal components. Santos (2005) parameterized the sensitivity coefficients and considered the effects of productive bienniality.

The thermal and hydric components are extremely important, because air temperature regulates the rate of vegetative and reproductive development, for example (PEREIRA et al., 2008). The ideal range of mean annual temperature for coffee is 18-21 °C, according to DaMatta (2007), and the development and ripening of the fruits are accelerated above 23 °C, leading to the loss of quality (CAMARGO, 1985). A temperature of -2 °C near the leaves damages the tissues. This temperature generally corresponds to 1-2 °C measured in the standard meteorological shelters, depending on the topographic conditions of the plantations (DAMATTA, 2007). Water deficit is a main factor that reduces crop yield (CARVALHO et al., 2011; MARTINS et al., 2015), because it decreases the rate of evapotranspiration (AMARASINGHE et al., 2015). Water deficits are higher when water stress occurs after fertilization, thus inhibiting fruit filling (CAMARGO, 2010).

Agrometeorological models for estimating yield traditionally use climatic data from surface meteorological stations (SMS), which is a limitation; because Brazil does not have, a network of stations that meets all needs (PEREIRA et al., 2002). The Institute Nacional of Meteorology (INMET) has a surface station approximately every 12 220 km^2 in the coffee-producing areas, much less than needed.

Using grid meteorological data (GD) is one method to solve this problem. GDs are models that combine information from a variety of sources, e.g. surfaces, oceans and remote sensing, and estimates and predicts the weather by dividing Earth's surface and atmosphere into a 3-dimensional grid (Bechtold et al., 2008). The predictions of time at each grid point are made using nonlinear equations applied for optimizing the mass and energy balance, and the input vectors are the meteorological elements, mainly temperature, humidity, atmospheric pressure and wind. Examples of GDs include the INPE/CPTEC global atmospheric model, the Global Data Assimilation System used by the National Forecasting Center and the ECMWF, which is an intergovernmental organization, composed of 34 countries, and the NASA.

The ECMWF forecasting system consists of three components: a general circulation model (coupled to an ocean-wave model), a system for assimilating data and, since 1992, a set forecasting system. The TL511L60 (ECMWF) model forecasts up to 10 days and is currently testing seasonal forecasts (six months). The ECMWF has a grid of 8 300 760 points in the three dimensions, with 60 levels of approximately 64 km of altitude, nearly the entire troposphere, and a horizontal grid of 0.25°, corresponding to an area of approximately 625 km^2 (25 × 25 km). The meteorological variables at each of these grid points are recalculated every 20 minutes up to ten days ahead. The total number of calculations is about 52 × 10^{12}, which takes approximately 95 minutes for the forecast with the current Fuijitsu VPP 700 (PERSSON, 2001).

The NASA-POWER project was developed to provide meteorological information for direct use for architecture, power generation and agrometeorology at a horizontal resolution of 1 × 1°, corresponding to an area of approximately 12 347 km^2 (111.12 × 111.12 km) at the equator. NASA-POWER compiles information from various data sources, directly and those derived from GDs. For example, real-time daily data for air temperature and relative humidity are obtained from the system of the Global Model and Assimilation Office (GEOS-4), and precipitation data are obtained from the Global Precipitation Climate Project.

The use of these meteorological data in grids has gained prominence in agricultural studies in recent years (GASPAR et al., 2015), such as the modeling of wheat phenology in the USA (WHITE et al., 2008), the simulation of the production potential of maize in China (BAI et al., 2010) and the simulation of the water satisfaction index for maize in South Africa (MOELETSI, WALKER, 2012). The GD data solve some limitations of surface data, such as uncertainties in measurement errors, e.g. connection errors in the data logger, clogged rain sensors, broken cables and mainly the lack of calibration of sensors in general (MORAES et al., 2014); missing data in climatic series, preventing reliable conclusions about the management of natural resources (MOELETSI; WALKER, 2012) and the time lag between collecting and receiving data (MORAES et al., 2012).

For Brazil, ECMWF and NASA-POWER have approximately 13 626 and 690 grid points, respectively, in contrast to 260 for the INMET SMS. The objective of this study was to estimate coffee yield in the coffee-producing regions of São Paulo and Minas Gerais using the model by Santos and Camargo (2006), based on meteorological data from GDs, ECMWF and NASA.

3.2 MATERIAL AND METHODS

Data for maximum, minimum and mean surface air temperatures were obtained for September 1992 to December 2015, and precipitation data were obtained for June 2000 to January 2016, for all locations (Table 1). These data were organized on a scale of ten-day period (10 days) as described by Santos and Camargo (2006). Data from 13 locations (Figure 1) were important for coffee production in Brazil (IBGE 2016), seven in the state of Minas Gerais (MG) and six in the state of São Paulo (SP) (Table 1). In the calibration stage of the coefficients of the model developed by Santos and Camargo (2006), 10 locations and 3 locations were used for the independent data test. Daily minimum, mean and maximum air temperatures and precipitation were provided by regional agricultural cooperatives for the locations in MG and by the Agronomic Institute for the SP locations.

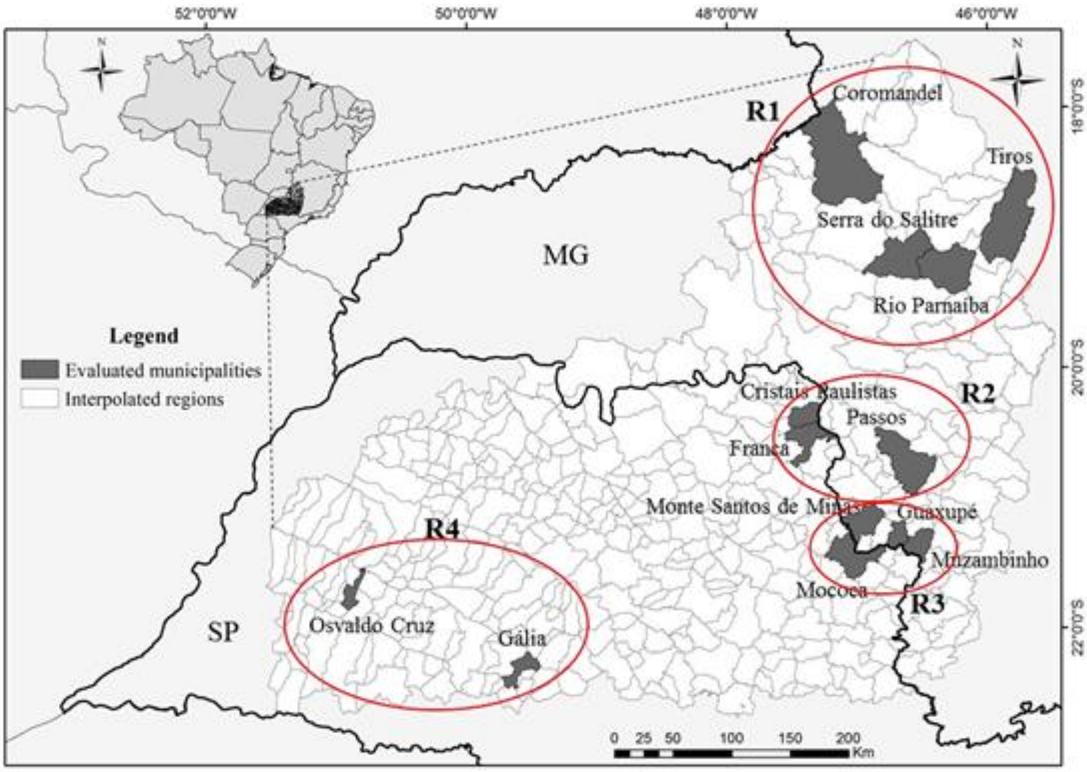

Figure 1. Locations studied in southeastern Brazil. Legend: R1, Cerrado Mineiro Region; R2, Alta Mogiana; R3, Sul de Minas and R4, Noroeste Paulista.

Data for coffee yields (in sacks of 60 kg ha^{-1}) for 2001-2014 were provided by regional agricultural cooperatives for the MG locations and by the Institute of Agricultural Economics for the SP locations.

Cerrado Mineiro (Coromandel, Tiros and Serra do Salitre) has a humid mesothermic climate based on the climatic classification of Thornthwaite (1948), with moderate water deficits in winter and with relative evapotranspiration <48 mm. The climate is similar in the other regions of MG and SP, but with or without small water deficits. Only Osvaldo Cruz has a subhumid climate. Serra do Salitre has the highest altitude at 1244 m, followed by Franca at 1026 m (Table 1).

Table 1. Altitude, climatic classification and observed productivities of the studied sites.

Locations	Altitude (m)	Climate[1]	Yield (sc ha $^{-1}$)[2]	Period	Calibration/Test
			CERRADO MINEIRO		
Coromandel-MG	928	$B'_2wB'_3a'$	37.56 ± 8.59	2005/2014	Calibração
Tiros-MG	977	$B'_2wB'_3a'$	34.38 ± 8.97	2005/2014	Calibração
Serra do Salitre-MG	1244	$B'_2wB'_3a'$	48.11 ± 19.18	2005/2014	Calibração
Rio Paranaíba-MG	1073	$B'_2wB'_3a'$	26.90 ± 6.30	2007/2015	Teste
			SUL DE MINAS GERAIS		
Guaxupé-MG	824	$B'_2rB'_4a'$	42.01 ± 14.56	2005/2014	Calibração
Muzambinho-MG	1100	$B'_2rB'_4a'$	18.80 ± 6.40	2007/2015	Teste
Monte Santo de Minas-MG	892	$B_3rB'_3a'$	39.24 ± 17.14	2002/2014	Calibração
Mococa-SP	665	$B'_2rB'_4a'$	19.13 ± 3.20	2002/2014	Calibração
			ALTA MOGIANA		
Franca-SP	1026	$B'_3rB'_3a'$	20.69 ± 3.14	2002/2014	Calibração
Passos-MG	745	$B'_3rB'_3a'$	22.20 ± 7.39	2007/2015	Teste
Cristais Paulista –SP	996	$B'_3rB'_3a'$	19.29 ± 4.42	2002/2014	Calibração
			NOROESTE PAULISTA		
Gália –SP	522	$B'_1rB'_4a'$	18.75 ± 2.53	2002/2014	Calibração
Osvaldo Cruz-SP	440	C_2rAa'	14.14 ± 1.85	2002/2014	Calibração

[1] Following the Thornthwaite climatic classification (1948). [2] Observed yield ± standard deviation.

The monthly potential evapotranspiration (ETP) was calculated using the model by Camargo (1971). Real evapotranspiration (ETR) and water deficit (DEF) were estimated by the sequential water-balance model developed by Thornthwaite and Mather (1955), using an available soil-water capacity equivalent to 100 mm on the sequential scale proposed by Rolim et al. (1998). Estimates of ETP take only meteorological factors into account, and estimates of ETR also include the soil-water conditions and DEFs due to the interaction of the soil-plant-atmosphere system (Equation 1).

$$DEF = ETP - ETR \tag{1}$$

3.2.1 ECMWF data

The ten-day data for minimum mean and maximum air temperatures and precipitation were collected from European Center for Medium-Range Weather Forecasts. These data are freely distributed on the internet and are available at http://spirits.jrc.ec.europa.eu/. These data are available in raster format, requiring extraction by software manipulation of satellite images, ENVI. Spatialization used QGIS software for characterizing the locations.

3.2.2 NASA data

The data series for daily average air temperature and precipitation were downloaded from NASA-POWER (Stackhouse 2010) (http://power.larc.nasa.gov), which allows downloading time series in ASCII format for a given geographical coordinate chosen by the user.

3.2.3 Santos and Camargo (2006) model

The model for estimating coffee yield developed by Santos and Camargo (2006) (Y_{SC}) is based on an agrometeorological model for estimating the loss in yield (Q (%)) proposed by Camargo et al. (2003), a mathematical model containing phenological and hydric (DH) components and thermal penalization factors, as:

$$Q\ (\%) = fDH\ fTMIN\ fTMAX \qquad (2)$$

where Q (%) is the relative loss of coffee yield and fDH, fTMIN and fTMAX are the penalization factors for water deficiency, minimum temperature and maximum temperature, respectively.

The phenological component model is used only to estimate the beginning of the full-flowering phase of the coffee plants that occurs after an accumulation of potential evapotranspiration of 350 mm, according to Camargo et al. (2003). This accumulation begins in January of the first phenological year (Figure 2).

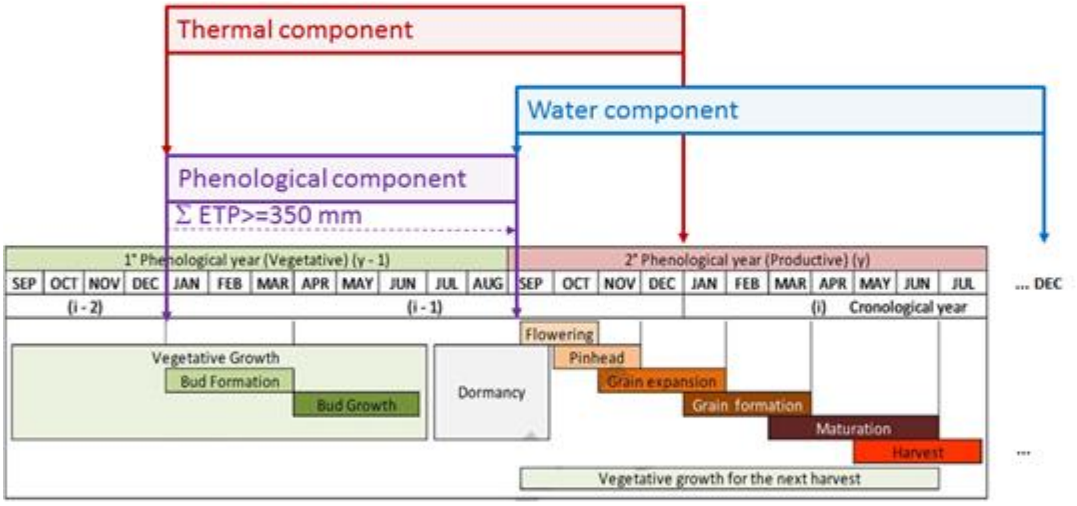

Figure 2. Schematic representation of Santos and Camargo model (2006). Modified from Aparecido et al. (2016).

The hydric component is based on the sequential water balance from the beginning of flowering (estimated) to the 36th ten-day period, where water deficit is quantified for the ten-day productivities of the adjusted [1 - (ETr/ETp)] ratios (Ky) for the four phenological stages (floral induction, flowering, granulation and maturation) (Equation 3) analyzed by different coefficients of sensitivity of the crop to the water deficit (ky). This component is similar to those proposed by Doorenbos and Kassan (1979) and Picini (1998) that consider the penalties for ten-day sums.

$$fDH = \prod_{i=1}^{n}\left[1-\left(ky\left(1-\frac{ETR}{ETP}\right)\right)\right] \quad (3)$$

Where: Π is the output and ky is the coefficient of sensitivity to water deficiency, i is the decay.

The thermal component of the model is based on minimum and maximum adverse temperatures in the year prior to production, from January to December, and which interfere with crop yield. We used the Gaussian penalization model developed by Camargo et al. (2003) for the absolute minimum temperatures, which estimates the percentage (%) of losses from annual absolute minimum temperatures <2 °C, obtained in meteorological protection, for the year prior to production as:

$$fTMIN = 1-\left[a_1\, exp^{\frac{-(x1-b1)^2}{2\,C1^2}}\right] \quad (4)$$

where fTMIN is the percentage of loss based on annual absolute minimum temperature, X1 is the absolute minimum shelter temperature of the year prior to production (°C), a_1 (140.3445) is a coefficient for the magnitude of the minimum temperature penalization, b1 (-2.9580) is a coefficient representing the temperature (°C) at the highest penalization and C1 (1.7701) is a coefficient representing the sensitivity of the model at low temperatures (magnitude of the penalization on the x-axis) (δmassa δ°C^{-1}).

We used the Gompertz model developed by Camargo et al. (2003) for high temperatures during flowering, which considers the penalization for absolute maximum air temperatures >34 °C during the first three 10-day periods after full anthesis, 11, 12 and 13 °C. The value of 34 °C can be replaced by the average temperature of 23 °C:

$$fTMAX = 1 - \left[a_2\, exp^{-exp^{(b2-(c2\, x2))}}\right] \tag{5}$$

where fTMAX is the penalization by mean ten-day temperatures during the flowering and pelleting stages (%), X2 is the average temperature (°C) of the first three ten-day periods after full bloom, a_2 (141.77) is a coefficient for the maximum coffee penalization for high temperatures, b2 (17.9486) is a reference coefficient for the beginning of the penalization and C2 (0.6782) is a coefficient for the sensitivity of the model at elevated temperatures.

Santos and Camargo (2006) incorporated another penalization factor (ky0) for the previous year's yield (Yaa) and potential yield (Yp). The values of ky0 tested varied from 0.60 (lower influence of Yaa) to 1.05 (higher influence of Yaa). The general agrometeorological model for estimating coffee yield based on the penalization functions can thus be expressed by:

$$Yest = \left[ky0\left(\frac{Yaa}{Yp}\right)\right]\left[1-\left(ky\left(1-\frac{ETr}{ETp}\right)\right)\right][1-fgeada]\,[1-fTMAX] \tag{6}$$

where Yest is the estimated yield (kg ha^{-1}) following Santos and Camargo (2006), Yaa is the previous year's yield (kg ha^{-1}) and Yp is the potential yield of the crop (kg ha^{-1}), which is the highest yield in the historical series for the location plus 10%. Ky0 corresponds to the coefficient of penalization relative to the previous year's yield and function of the hierarchical level considered (farm, rural property and location)

(PICINI et al., 1999; CARVALHO et al., 2003). The Ky0 that best represented the hierarchical level of rural property was 0.63 (Santos, 2005).

3.2.4 Parameterizations of the Santos and Camargo (2006) model

The equations used to quantify the adverse thermal penalties (fTMIN and fTMAX) verified that a1 and a_2 represented the amplitude of the curve, i.e. the maximum penalization (%) as a function of adverse temperatures (Figure 3). These values should therefore not be equal to those of approximately 140% used by Santos and Camargo (2006). The maximum penalty should be 100%.

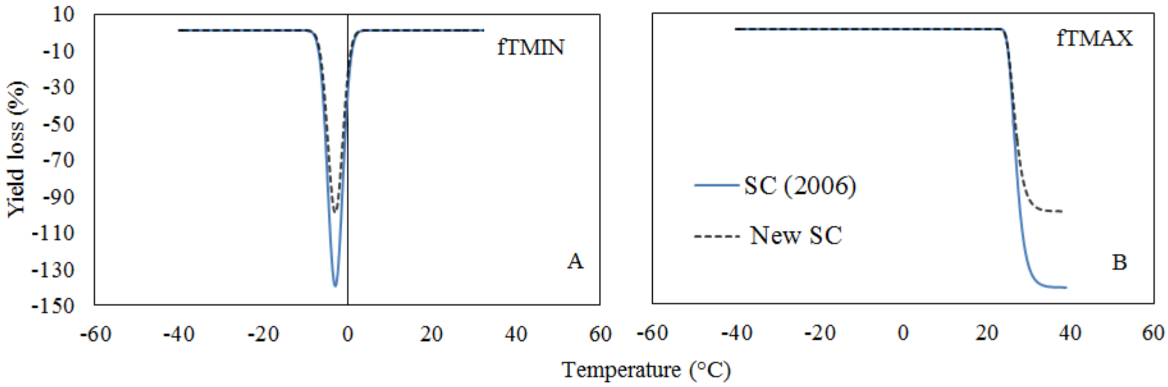

Figure 3. Verification of the coefficients used in the equations of penalties by minimum (A) and maximum (B) temperatures. Legend: SC as proposed by Santos and Camargo (2006) and new SC new calibration.

The coefficients a1 and a_2 were set at 100% for the location hierarchical level, and the others were calibrated, including the coefficients of sensitivity for the yield of the previous year (Ky0) and the water deficit (ky). Each location was calibrated using the generalized reduced gradient 2 method proposed by Lasdon et al. (1978), which is considered the most efficient for the general solution of nonlinear-optimization problems. The calibrations simultaneously adjusted several parameters of the model (ten-day ky, b1, c1, b2, c2 and ky0) to minimize the square of the errors between observed and estimated yield data.

After calibration with the surface data, these coefficients were used in the models with meteorological input data from ECMWF and NASA. A new calibration with the ECMWF and NASA data attempted to improve the performance of the models (Table 2). The applicability of the model will be better if only grid data are used, because the spatial representation will be greater, but it is necessary to verify if the necessary calibrations are sufficiently robust compared to the actual calibrations of the surface.

Table 2. Analyzed models, calibrations performed and source of applied data for yield estimates.

Analyzed model	Calibration in relation to to the data of:			Estimation of yield using data:		
	Superfície	ECMWF	NASA	Superfície	ECMWF	NASA
Y_{SC-S}	✖			✖		
$Y_{ECMWF-S}$	✖				✖	
Y_{ECMWF}		✖			✖	
Y_{NASA-S}	✖					✖
Y_{NASA}			✖			✖

where Y_{SC-S} is the yield-estimation model developed by Santos and Camargo (2006), parameterized with the surface data, Y_{ECMWF} is the data estimated by the Santos and Camargo model using the climatic data of the ECMWF systems , which used two calibrations, one with the surface data ($Y_{ECMWF-S}$) and then one with ECMWF data (Y_{ECMWF}) and Y_{NASA} is the data estimated by the model of Santos and Camargo (2006) using the data of the NASA systems, which used two calibrations, one with the surface data (Y_{NASA-S}) and then one with the same data as the NASA model (Y_{NASA}).

An equation for estimating the final yield was another adjustment to the Santos and Camargo (2006) model. The yield (kg ha^{-1}) is estimated by multiplying four penalties (yield of the previous year, fDH, fTMIN and fTMAX) (Equation 6), but the output of this equation is a fractional value. Yest in the original equation is thus equal to the relative breakage percentage (Q%). The percentage of relative yield loss should be subtracted from potential yield for estimating yield in kg ha^{-1} or sacks ha^{-1}.

The penalization of the previous year associated with yield, [ky (Yaa/Yp)], was also changed. We must subtract the result from 1, because the value obtained in this relationship weighted by ky is how much "left" of the break and not how much it "broke", so the relationship becomes 1 - [Ky0 (Yaa/Yp)]. The percentage of decrease in relative productivity therefore becomes Equation 7.

$$Q\% = \left[1 - \left(ky0\left(\frac{Yaa}{Yp}\right)\right)\right] [fDH]\,[fTMIN]\,[fTMAX] \quad (7)$$

The model for estimating coffee yield in sacks ha^{-1} parameterized by the cited observations is:

$$Yest = [(1 - Q\%)\,Yp] \quad (8)$$

where Yest is the estimated yield in sacks ha^{-1} for the suggested modifications.

The models with independent data were tested for three locations (Table 1). Geographical proximity was the criterion used for choosing the model to be applied in each region. The model for Muzambinho thus used the models calibrated for Guaxupé, the model for Passos used those calibrated for Franca and the model for Rio Paranaíba used those calibrated for Serra do Salitre. These three locations are approximately 21.4, 83 and 43.7 km from the calibrated locations, respectively.

3.2.5 Statistical analysis nálises estatísticas

The models were evaluated by comparing the estimated and observed productivities. The models calibrated for ECMWF and NASA were evaluated by comparing the parameterized productivities for Santos and Camargo (2006) model.

The accuracy, tendency and precision of the models were determined using the mean absolute percentage error (MAPE), systematic root mean square error (RMSEs) and non-systematic root mean square error (RMSEn), respectively.

$$MAPE = \frac{\sum_{i=1}^{N}\left(\left|\frac{Yest_i - Yobs_i}{Yobs_i}\right| * 100\right)}{N} \quad (9)$$

$$RMSEs = \sqrt{\frac{\sum_{i=1}^{N}(Yobs_i - Yest_c)^2}{N}} \quad (10)$$

$$RMSEn = \sqrt{\frac{\sum_{i=1}^{N}(Yest_i - Yest_c)^2}{N}} \quad (11)$$

where Y_{obsi} is the ith yield observed in the field, Yesti is the yield estimated by the various models, YestC is the yield estimated by linear fit and N is the number of datapoints. The MAPEs for the construction of maps for the different sources of climatic data (ECMWF and NASA) were used for general extrapolations of the results, interpolating the values by inverse distance weighting.

3.3 RESULTS AND DISCUSSION

The NASA data for minimum temperatures were most similar to the surface data, but the ECMWF data for maximum temperatures were most similar to the surface data. The water deficit did not differ significantly between the data sources (Figure 4).

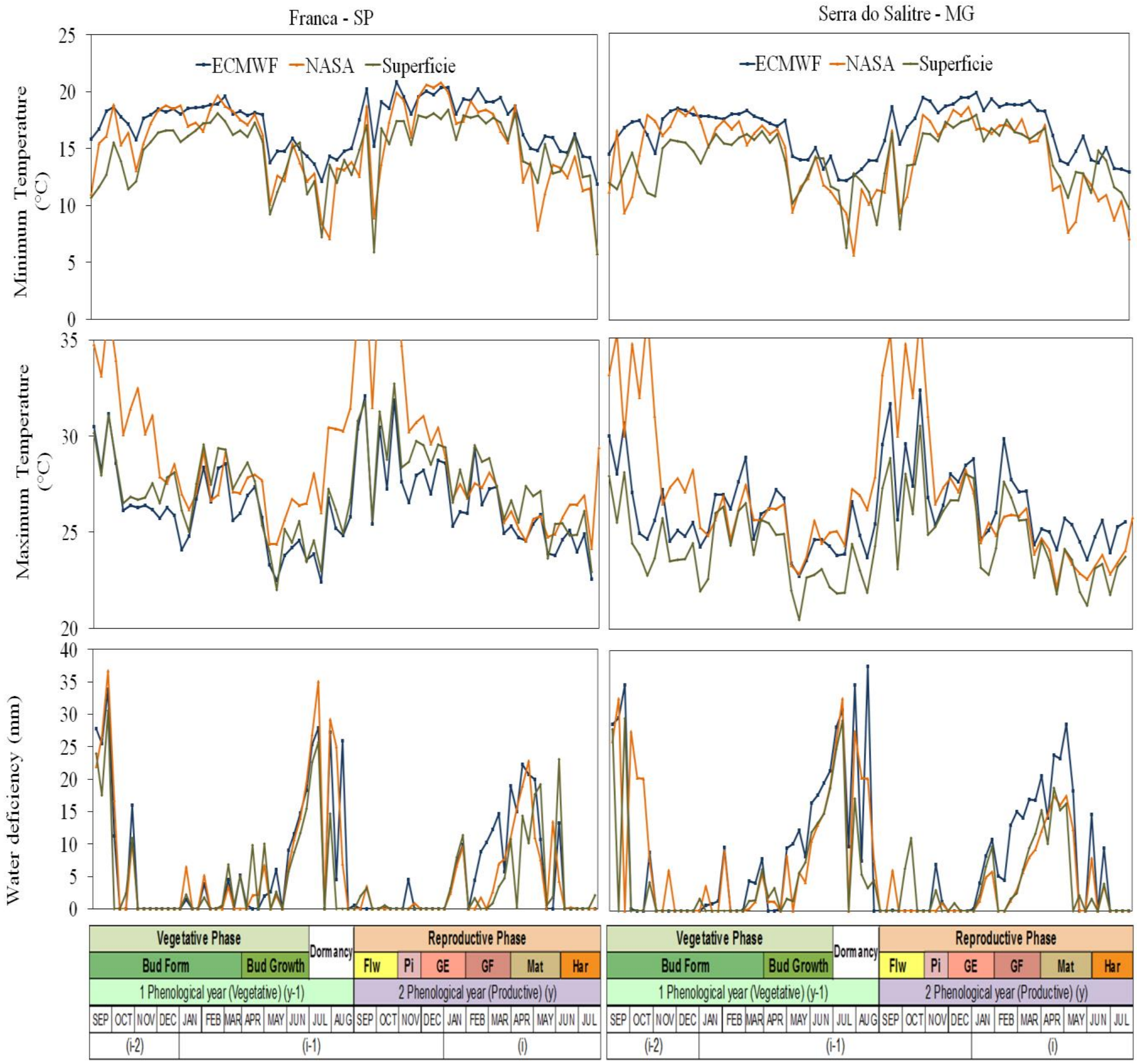

Figure 4. Minimum, maximum and water deficit data for the years 2011, 2012 and 2013 from surface, ECMWF and NASA. Legend: Flw, flowering; Pi, Chumbinho; GE, Fruit expansion; GF, Granulation of fruits; Mat, Maturation; Har, Harvest.

Coffee yield has a biennial characteristic in the coffee-producing regions in Brazil (Figure 5). This bienniality is due to the relationship between plant phenology and climate. The plant applies energy to photoassimilates in the first phenological year for vegetative development, which decreases yield. The plant, with its developed canopy, then applies energy in the second phenological year to the reproductive phase, which increases yield. This relationship has been described in more detail by Aparecido et al. (2016).

The Santos and Camargo (2006) model can be simplified by adding 10% to the maximum yield in a region to define the potential yield (Figure 5). Potential yield depends only on solar irradiance, air temperature and genetic potential. Wang et al. (2015) determined that the potential yield of coffee in Uganda may be 70-80% higher than the average actual yield. This simplification affects the calibrations of the coefficients, because they must compensate for this presupposition. We decided to use this simplification, however, due to the lack of data for potential yield in MG and SP. The model is extensively used for coffee-producing areas, so this simplification helps us to understand how the use of grid data affects the estimates made by the model compared to those originally made from surface data.

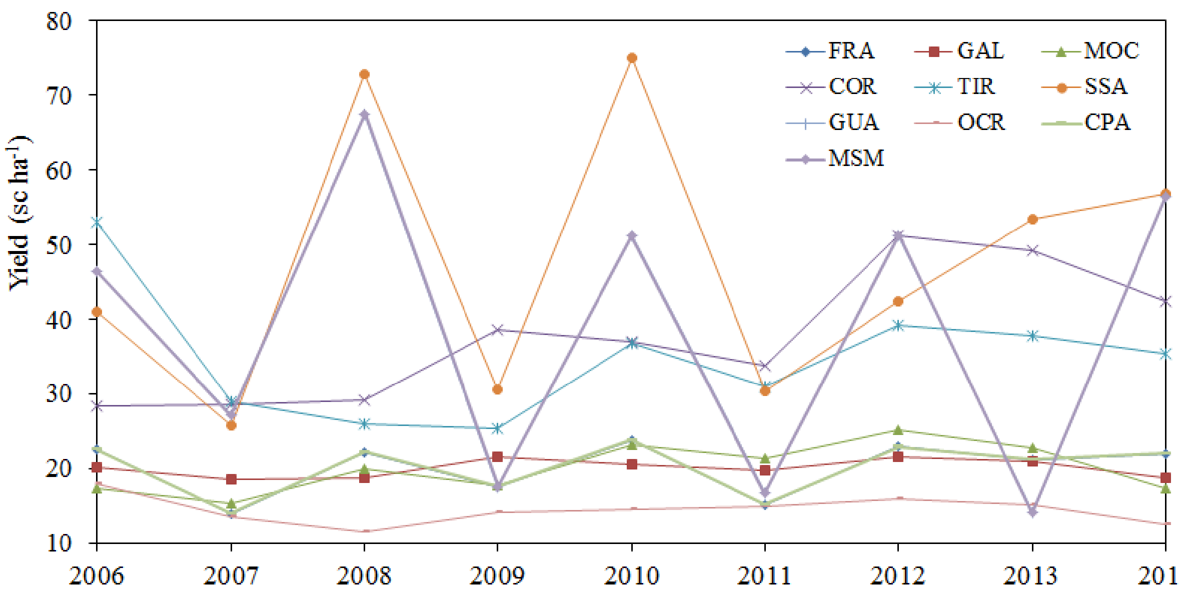

Figure 5. Yield observed in coffee plantations in the regions analyzed between 2006 and 2014. The arrows indicate the maximum yield observed in the period in each locality. Legend: Localities of Franca (FRA), Coromandel (COR), Guaxupé (GUA),

Monte Santo de Minas (MSM), Gália (GAL), Tiros (TIR), Osvaldo Cruz (OCR), Mococa (MOC), Serra do Salitre (SSA) and Cristais Paulista (CPA).

We calibrated the regional genetic coefficients for three conditions, the first using surface meteorological data and the second and third using the ECMWF and NASA data, respectively. Data for the entire year were used for each location. The coefficients did not notably differ between the calibrations (Table 3). Franca had higher air temperatures, for a lower value of b1 in the calibration with the surface data of -5.9 °C. This temperature was approximately -2 °C for the GD data. Serra do Salitre had the highest b1 for the NASA data (2 °C), indicating that the penalization was maximal at this temperature.

Agronomic frost is defined as the occurrence of minimum temperatures that kill plant tissues or the entire plant, thereby affecting yield (CAMARGO, 1993). Vasconcellos et al. (2013) estimated that 100% of the decreases in yield in coffee plantations in 1990 and 1994 were due to the occurrence of absolute minimum temperatures of -2.8 °C. A minimum temperature of 2 °C is the limit below which the yield of coffee trees begins to decrease (Camargo et al., 1993). The temperature that actually defines the level of damage, however, is the minimum observed on the grass, which can be up to 4 °C lower than that measured by sensors at 1.5 m. BOOTSMA (1976) and VASCONCELLOS et al. (2013) reported that temperatures <-2 °C, which in the model caused total breakage, did not always cause complete crop failure. This difference between results in years with frost may have been due to the way the frost affected the coffee stand; the model considers a total loss of the harvest, i.e. the branches and leaves are affected by the cold, whereas the frost may have only affected the outside of the tree.

Only Gália, Guaxupé and Cristais Paulista recorded temperatures <2 °C from 2006 to 2014, in August 2008, June 2009 and June 2011 in Gália, June and August 2011 and July 2013 in Guaxupé and June 2009 in Cristais Paulista.

The coefficient b2 indicates a reference temperature for the start of penalization due to maximum temperatures in the first three ten-day periods after full bloom. The penalization is a function of the maximum temperature, but the average temperature is entered into the model, because average temperatures >23 °C are equivalent to the maximum temperature of 34 °C (Camargo et al., 2003). Temperatures >23 °C can accelerate the development and maturation of the fruit, leading to a loss of quality (CAMARGO, 1985; DAMATTA et al., 2008).

The calibrated values of b2 changed little, near 18 °C, only in Osvaldo Cruz, Coromandel and Tiros that had upper values of 25 °C for the calibration with the NASA data. Briefly, b1 and b2 tended to be overestimated in the calibration of the genetic coefficients with the NASA data, likely due to overestimations of maximum temperatures by NASA for the locations (Figure 4).

The coefficients c1 and c2 represent the model sensitivities at the minimum and maximum temperatures, respectively. The model is more sensitive at lower values of c1 and c2. Mean c1 and c2 for all locations were 1.86 and 0.63, respectively. The mean c2were lowest in Osvaldo Cruz (0.36) in the calibration with the ECMWF data. These values are similar to those reported by Santos and Camargo (2006).

The lower the coefficient (ky0), the higher the previous year's yield penalty (SANTOS, 2005). Guaxupé had the highest penalty, with a ky0 of 0.13 for the calibration with the ECMWF data. Gália had a lower effect on yield in the previous year, with a ky0 of 1. Santos and Camargo (2006) concluded that the best ky0 was 0.96, but for the hierarchical level. A ky0 of 0.45 is recommended for the municipality hierarchical level (Camargo et al., 2005), indicating that the larger the scale (municipality), the lower the amplitude of the productivities.

Table 3. Genetic coefficients of the model Santos and Camargo (2006) calibrated from different sources of meteorological data.

	Locations									
	COR	TIR	SSA	GUA	MSM	MOC	FRA	CPA	GAL	OCR	Média ± Desvio Padrão
Coeficientes genéticos regionais calibrados com dados de superfície (Y_{SC-S})											
a1	100	100	100	100	100	100	100	100	100	100	100 ± 0
b1	0.61	-3.00	0.02	-2.96	-2.96	-2.88	-5.91	-1.43	1.08	-2.96	-2.04 ± 2.12
c1	2.37	2.62	1.91	1.77	1.77	1.87	3.55	1.10	0.33	0.26	1.75 ± 1
a2	100	100	100	100	100	100	100	100	100	100	100 ± 0
b2	18.85	18.23	17.95	17.95	18.38	17.74	18.31	18.56	18.02	21.65	18.56 ± 1.13
c2	0.65	0.67	0.73	0.68	0.66	0.68	0.67	0.66	0.68	0.55	0.66 ± 0.04
Ky0	0.63	0.42	0.79	0.31	0.20	0.70	0.21	0.63	1.00	0.56	0.54 ± 0.25
Coeficientes genéticos regionais calibrados com dados do modelo ECMWF (Y_{ECMWF})											
a1	100	100	100	100	100	100	100	100	100	100	100 ± 0
b1	1.66	-2.96	-2.96	-4.32	-2.96	-2.96	-2.96	-2.96	-2.96	-2.96	-2.63 ± 1.56
c1	2.58	1.77	1.77	0.78	1.77	1.77	1.77	1.77	1.77	1.77	1.75 ± 0.42
a2	100	100	100	100	100	100	100	100	100	100	100 ± 0
b2	18.85	18.27	17.95	17.95	17.95	18.81	18.39	17.03	17.99	25.21	18.84 ± 2.29
c2	0.65	0.67	0.68	0.68	0.68	0.65	0.73	0.73	0.68	0.36	0.65 ± 0.10
Ky0	0.56	0.50	0.54	0.13	0.15	0.63	0.61	0.20	0.92	0.68	0.49 ± 0.25
Coeficientes genéticos regionais calibrados com dados do modelo NASA (Y_{NASA})											
a1	100	100	100	100	100	100	100	100	100	100	100 ± 0
b1	-2.00	-1.00	2.05	-3.00	-2.00	-3.46	-2.00	1.01	-2.00	-2.97	-1.54 ± 1.77
c1	3.28	1.80	1.10	2.33	2.23	2.49	2.44	1.59	1.84	1.76	2.09 ± 0.60
a2	100	100	100	100	100	100	100	100	100	100	100 ± 0
b2	25.00	25.00	18.01	17.95	17.32	20.00	18.31	21.11	20.00	25.00	20.77 ± 3.13
c2	0.44	0.44	0.68	0.68	0.47	0.79	0.67	0.58	0.61	0.43	0.58 ± 0.12
Ky0	0.63	0.63	0.54	0.63	0.34	0.63	0.28	0.18	0.63	0.68	0.52 ± 0.18

Legend: a_1 = indicates the amplitude of the penalization by minimum temperature (%); b1 = is the temperature in (° C) where the highest penalization occurs; c1 represents the sensitivity of the model at low temperatures; $a_2$indicates the maximum coffee penalization for high temperatures; b2 is the reference coefficient where the beginning of the penalization begins; c2 indicates the sensitivity of the model at elevated temperatures, the higher the value of the faster coefficient it starts to penalize. COR- Coromandel; TIR- Tiros; SSA- Serra do Salitre; GUA- Guaxupé; MSM- Monte Santo de Minas; MOC- Mococa; FRA- Franca; CPA- Cristais Paulista; GAL-Gália; ORC- Osvaldo Cruz.

The hydric penalization component considers the stress of the crop due to water deficiency, weighted by ky, multiplicatively in the various phenological stages (SANTOS; CAMARGO, 2006). Santos (2005) tested several pre-established sequences of ky. We, however, calibrated these values for each location and model. ky varied from 0 to 1 for the ten-day period (Figure 6). ky is correlated with the slope of the ratio Y/Yp = ky ETR/ETP so should not vary among locations (Doorenbos and Kassam, 1979). We performed the calibrations for the locations, however, to verify the robustness of these ten-day coefficients and found that the recalibrated values were very similar to those suggested by Santos and Camargo (2006).

A value of ky of 1, as in Franca, indicates that a decrease of, for example, 10% in the deficit of relative evapotranspiration equals a 10% decrease in yield, indicating a very sensitive site for water deficiency. The locations with low values of ky, such as Serra do Salitre, though, were less sensitive, probably because these locations had lower water deficits; the largest variations were associated with the attainable yield due to water availability. For example ky was 1 in Franca in 12th ten-day but was 0.04 in Serra do Salitre. The accumulated DEF up to this ten-day period in 2011 was 120 mm in Franca and 70 mm in Serra do Salitre.

The highest values of ky were associated with the phenological stages of floral induction, flowering and granulation, in agreement with the studies by Picini et al. (1999) and Santos and Camargo (2006). During the period between the end of rest in August, 22, 23 and 24th ten-days, and the beginning of full bloom, which normally occurs from September to October, between 25 and 30th ten-days, the Santos and Camargo (2006) model considers a fixed value of 0.02, similar to those in the present analysis.

We noted a lack of spatial and temporal compensation between the NASA and ECMWF data; ky calibrated with the NASA data was more similar to the values suggested by Santos and Camargo (2006). The NASA data have a grid sixteen times larger than that of the ECMWF data and thus have lower spatial detail. The NASA data, however have a better daily time resolution (10 days) than the ECMWF data. NASA's best time scale was thus generally more important than the spatial ECMWF scale for crop sensitivity to water deficiency (ky).

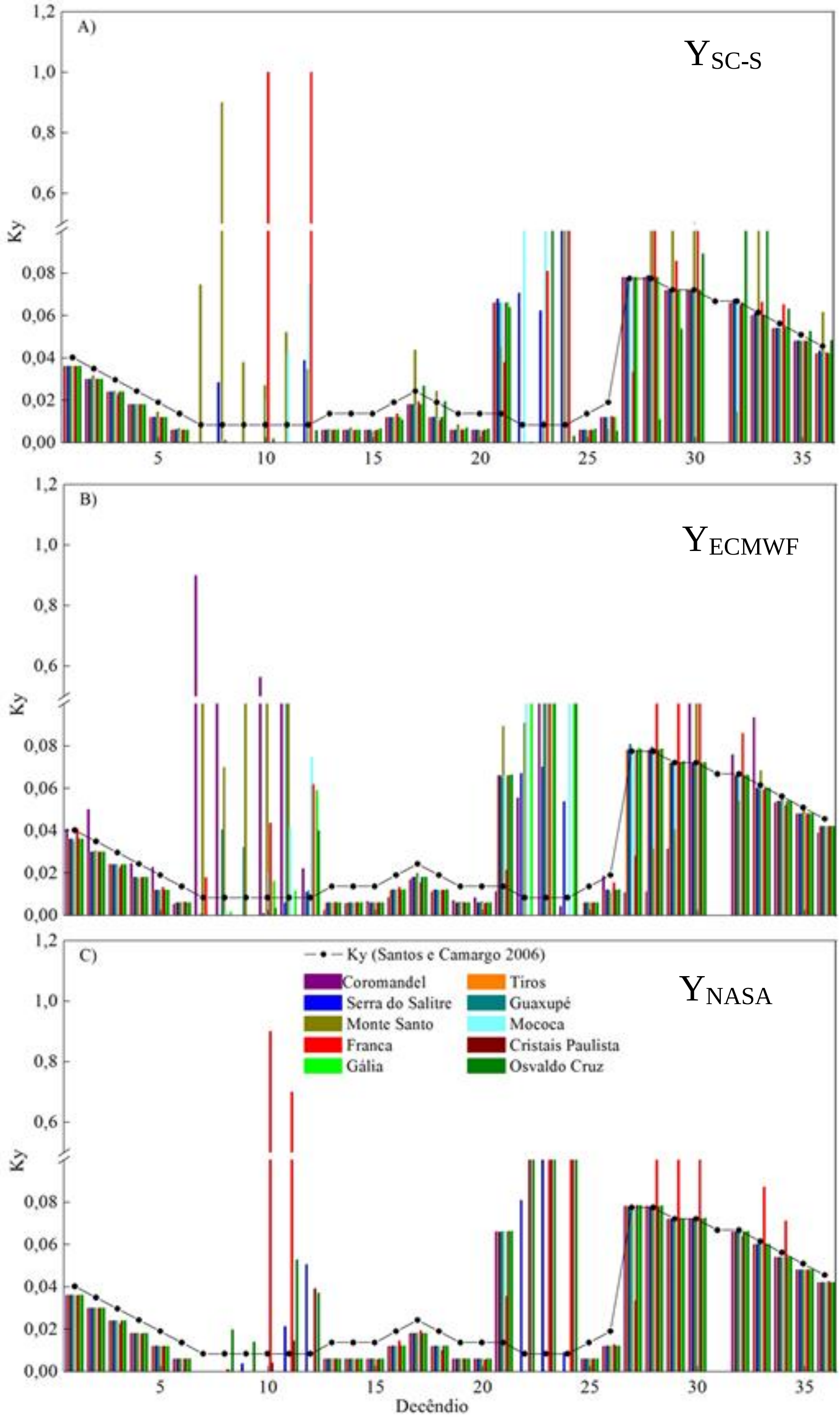

Figure 6. Values of water deficit sensitivity coefficient (ky) for the sites studied on a decendial scale for Surface (A), ECMWF (B) and NASA (C).

The yields estimated by combinations of calibrations with the different data sources were similar, and the accuracies of the models were generally satisfactory. We evaluated the performance of the models by comparing their estimated yields to the yields observed in the field (Y_{OBS}) and to the yields estimated by Santos and Camargo (2006) (Y_{SC-S}).

Mean MAPEs for Y_{OBS} were 23.19 and 23.76% for the Y_{ECMWF} and $Y_{ECMWF-S}$ models, respectively. The most accurate location was Gália, with a MAPE of approximately 10%. Monte Santo de Minas had the highest MAPE of approximately 58%, indicating that the models for this location could not accurately estimate yield (Table 4). The NASA models (Y_{NASA} and Y_{NASA-S}) performed similarly, with mean MAPEs of 24.61 and 27.98%, respectively. Gália also had the lowest MAPE, but Serra do Salitre had the lowest MAPEs for the ECMWF models, averaging 49% (Table 4).

The models calibrated with grid data (Y_{ECMWF} and Y_{NASA}) were more accurate than the models calibrated with surface data ($Y_{ECMWF-S}$ and Y_{NASA-S}), indicating that the uncertainties increased when data from different sources were combined. We also evaluated the performance of the grid-data models relative to the yields estimated by Santos and Camargo (2006) from surface data (Y_{SC-S}). The models using the ECMWF data were more accurate than those using the NASA data, with MAPEs of 9.37 and 20.31%, respectively, indicating that the spatial resolution of the ECMWF data was more important than the NASA temporal resolution in the final estimation of yield.

In summary, these results indicate that calibrating the model using surface data is not necessary and that the calibration with the ECMWF grid data can be used for accurately estimating yield by the Santos and Camargo (2006) model. Cerrado Mineiro had highest average MAPEs, followed by Sul de Minas, Alta Mogiana and Noroeste Paulista, with MAPEs of 27.1, 23.4, 14.1 and 10.5%, respectively.

Table 4. Accuracy of the models in relation to observed yield (Y_{OBS}) and yield estimated by the Santos and Camargo model, parameterized ($Y_{SC\text{-}S}$), using the mean absolute percentage error (MAPE).

Comparison	MAPE (%)								
about:	Y_{OBS}		Y_{SC-S}		Y_{OBS}		Y_{SC-S}		
Locais	$Y_{ECMWF-S}$	Y_{ECMWF}	$Y_{ECMWF-S}$	Y_{ECMWF}	Y_{NASA-S}	Y_{NASA}	Y_{NASA-S}	Y_{NASA}	Média ± Desvio Padrão
					Cerrado Mineiro				
COR	15.35	13.80	14.57	9.37	42.23	38.77	24.61	20.31	22.37 ± 12
TIR	23.36	24.44	17.33	17.63	37.73	37.12	30.28	28.76	27.08 ± 7
SSA	48.34	41.07	14.38	4.37	53.36	45.23	34.63	13.02	31 ± 18
					Sul de Minas				
GUA	18.89	28.58	24.14	34.98	25.99	25.33	29.3	26.45	26.07 ± 4
MSM	51.79	64.94	18.07	18.78	39.56	37.73	24.12	18.81	34.22 ± 17
MOC	14.54	13.73	5.79	4.85	11.42	8.53	7.45	8.68	9.37 ± 3
					Alta Mogiana				
FRA	22.28	13.84	15.12	10.45	13.6	10.81	7.02	3.35	12.05 ± 5
CPA	14.28	11.87	5.17	1.19	26.69	20.71	31.01	18.81	16.21 ± 10
					Noroeste paulista				
GAL	10.58	10.76	5.45	9.31	11.95	8.66	11	7.81	9.44 ± 2
OCR	12.48	14.55	6.67	6.59	17.25	13.25	17	6.14	11.74 ± 4
Média	23.18 ± 13	23.78 ± 16	12.66 ± 6	11.75 ± 9	27.97 ±13	24.61 ± 13	21.64 ± 9	15.24 ± 8	-

Legend: COR- Coromandel; TIR- Tiros; SSA- Serra do Salitre; GUA- Guaxupé; MSM- Monte Santo de Minas; MOC- Mococa; FRA- Franca; CPA- Cristais Paulista; GAL- Gália; ORC- Osvaldo Cruz.

The tendency of a model is the systematic distance between the estimated and observed values, so for good performance, the tendency of a model should be as low as possible, as it was for most locations in our study (Table 5). Gália had the lowest RMSE, 0.75 sacks ha^{-1}, taking into account the two comparisons (Y_{OBS} and $Y_{SC\text{-}S}$) for all calibrations. Serra do Salitre had the highest tendency with an average RMSE of 19.90 sacks ha^{-1}. Tendency values tend to increase, slightly overestimate low yields and underestimate the high yields as the hierarchical scale increases (plant, plot, municipality and region) (Camargo et al., 2005). RMSEs tended to be lower for the calibrations with the GD data (Y_{ECMWF} and Y_{NASA}) than with the surface data ($Y_{ECMWF\text{-}S}$ and $Y_{NASA\text{-}S}$). These results are similar to those of Santos and Camargo (2006), whose lowest and highest RMSEs for Mococa in SP were 5.4 and 12.4 sacks ha^{-1}, respectively.

Table 5. Trend of the models in relation to observed yield (Y_{OBS}) and yield estimated by the Santos and Camargo model, parameterized ($Y_{SC\text{-}S}$), through root mean square error systematic (RMSEs).

Comparison about:	RMSEs (sc ha^{-1})								
	Y_{OBS}		$Y_{SC\text{-}S}$		Y_{OBS}		$Y_{SC\text{-}S}$		
Locais	$Y_{ECMWF\text{-}S}$	Y_{ECMWF}	$Y_{ECMWF\text{-}S}$	Y_{ECMWF}	$Y_{NASA\text{-}S}$	Y_{NASA}	$Y_{NASA\text{-}S}$	Y_{NASA}	Média ± Desvio Padrão
				Cerrado Mineiro					
COR	8.79	5.34	9.14	4.77	13.94	13.41	9.68	6.56	8.9 ± 3
TIR	14.57	10.07	8.27	7.96	12.78	12.49	10.49	9.92	10.8 ± 2
SSA	30.18	20.64	7.68	0.99	24.36	19.65	48.49	7.19	19 ± 15
				Sul de Minas					
GUA	17.84	11.15	15.42	16.09	20.74	12.8	22.77	12.44	16.1 ± 4
MSM	27.93	18.91	2.33	5.09	15.12	12.35	9.46	8.01	12.4 ± 8
MOC	4.34	2.88	1.33	1.12	2.06	1.02	0.86	0.48	1.76 ± 1
				Alta Mogiana					
FRA	4.51	3.26	3.15	1.98	2.24	1.31	0.81	0.37	2.2 ± 1
CPA	1.52	1.13	1.45	0.18	6.47	3.64	7.9	4.19	3.31 ± 2
				Noroeste paulista					
GAL	2.16	2	1.22	1.87	1.96	1.14	1.78	1.78	0.7 ±0.3
OCR	3.05	2.12	0.22	0.19	3.21	1.92	2.38	0.44	1.69 ± 1
Média	11.49 ± 10	7.75 ± 7	5.02 ± 5	4.02 ± 5	10.28 ± 8	7.9 ± 6	11.46 ± 14	5.13 ± 4	-

Legend: COR- Coromandel; TIR- Tiros; SSA- Serra do Salitre; GUA- Guaxupé; MSM- Monte Santo de Minas; MOC- Mococa; FRA- Franca; CPA- Cristais Paulista; GAL- Gália; ORC- Osvaldo Cruz.

The precision of a model is an indication of the variation in a set of measurements: the higher the precision, the lower the variability between measurements (REVISTABW, 2014). The models had low RMSEn values, so they were precise (Table 6). Precision was highest for Osvaldo Cruz, Gália, Franca and Mococa, with an RMSEn near 1 sack ha^{-1}. This precision was higher than that of Santos and Camargo (2006), who found values between 2.4 and 9.3 sacks ha^{-1}. Precision also tended to be higher for the calibrated models (Y_{ECMWF} and Y_{NASA}) than those calibrated with surface data ($Y_{ECMWF\text{-}S}$ and $Y_{NASA\text{-}S}$).

Table 6. Precision of the models in relation to the observed yield (Y_{OBS}) and yield estimated by the Santos and Camargo model, parameterized (Y_{SC-S}), through the root mean square error non-systematic (RMSEn).

Comparison about:	RMSEn (sc ha^{-1})								
	Y_{OBS}		Y_{SC-S}		Y_{OBS}		Y_{SC-S}		
Locais	$Y_{ECMWF-S}$	Y_{ECMWF}	$Y_{ECMWF-S}$	Y_{ECMWF}	Y_{NASA-S}	Y_{NASA}	Y_{NASA-S}	Y_{NASA}	Média ± Desvio Padrão
				Cerrado Mineiro					
COR	5.59	1.83	5.53	3.31	10.00	8.83	7.82	7.73	6.33 ± 2
TIR	5.51	4.12	3.99	4.25	5.23	5.73	5.88	6.37	5.13 ± 0.9
SSA	7.74	5.76	2.27	2.37	7.67	5.99	30	0.31	7.76 ± 9
				Sul de Minas					
GUA	7.33	5.18	7.61	4.6	2.97	9.19	3.34	8.88	6.13 ± 2
MSM	16.06	7.46	8.97	6.65	5.6	6.19	5.53	7.12	7.94 ± 3
MOC	1.69	1.12	1	0.97	2.19	1.78	1.48	1.92	1.51 ± 0.4
				Alta Mogiana					
FRA	2.36	1.09	1.85	1.15	1.99	1	1.46	0.76	1.45 ± 0.5
CPA	3.76	2.43	3.21	0.33	0.73	2.18	1.36	2.12	2.01 ± 1
				Noroeste paulista					
GAL	2.25	0.97	0.95	0.7	2.39	1.97	1.77	1.5	1.56 ± 0.6
OCR	1.63	1.11	0.96	1.04	1.06	1	1.19	0.87	1.10 ± 0.2
Média	5.39 ± 4	3.10 ± 2	3.63 ± 3	2.53 ± 2	3.98 ± 3	4.38 ± 3	5.98 ± 8	3.75 ± 3	-

Legend: COR- Coromandel; TIR- Tiros; SSA- Serra do Salitre; GUA- Guaxupé; MSM- Monte Santo de Minas; MOC- Mococa; FRA- Franca; CPA- Cristais Paulista; GAL- Gália; ORC- Osvaldo Cruz.

The MAPEs decreased in the evaluation of the performance of the change of the models calibrated with surface data ($Y_{ECMWF-S}$ and Y_{NASA-S}) for the models fully calibrated with grid data (Y_{ECMWF} and Y_{NASA}), indicating an improvement in the response of the model when using only grid data for estimating the yield of coffee by the Santos and Camargo (2006) model. We therefore recommend using the ECMWF and NASA input data, with their calibrated coefficients, because the models performed better. Only Monte Santo de Minas and Guaxupé with the ECMWF data did not follow this variation (Figure 7). The possible reasons for low model accuracy (high MAPE) are low precision (low R^2 or high RMSEn) or high tendency (high RMSEs). Monte Santo de Minas and Guaxupé are at high altitudes and have mountainous relief, so these locations differed because of microclimatic factors such as relief.

Atmospheric conditions and topography are possible sources of error for the grid GD data (Che et al., 2016). White et al. (2008) reported larger differences between grid climatic data and terrestrial measurements in mountainous areas in the continental United States. All sources of surface and grid data have inherent errors and uncertainties. Uncertainties can be associated with the calibration of sensors, operation and human error or those associated with the operation of satellites, such as pixel size, sensor resolution, navigation time, algorithmic accuracy and geographical coincidence of the data captured instantaneously by a satellite with data measured in the field (CHE et al., 2016). Wu et al. (2005), Holder et al. (2006), Mahmood et al. (2006) and White et al. (2008) have reported that analyses of data sets from surface stations can have errors as high as the GD errors.

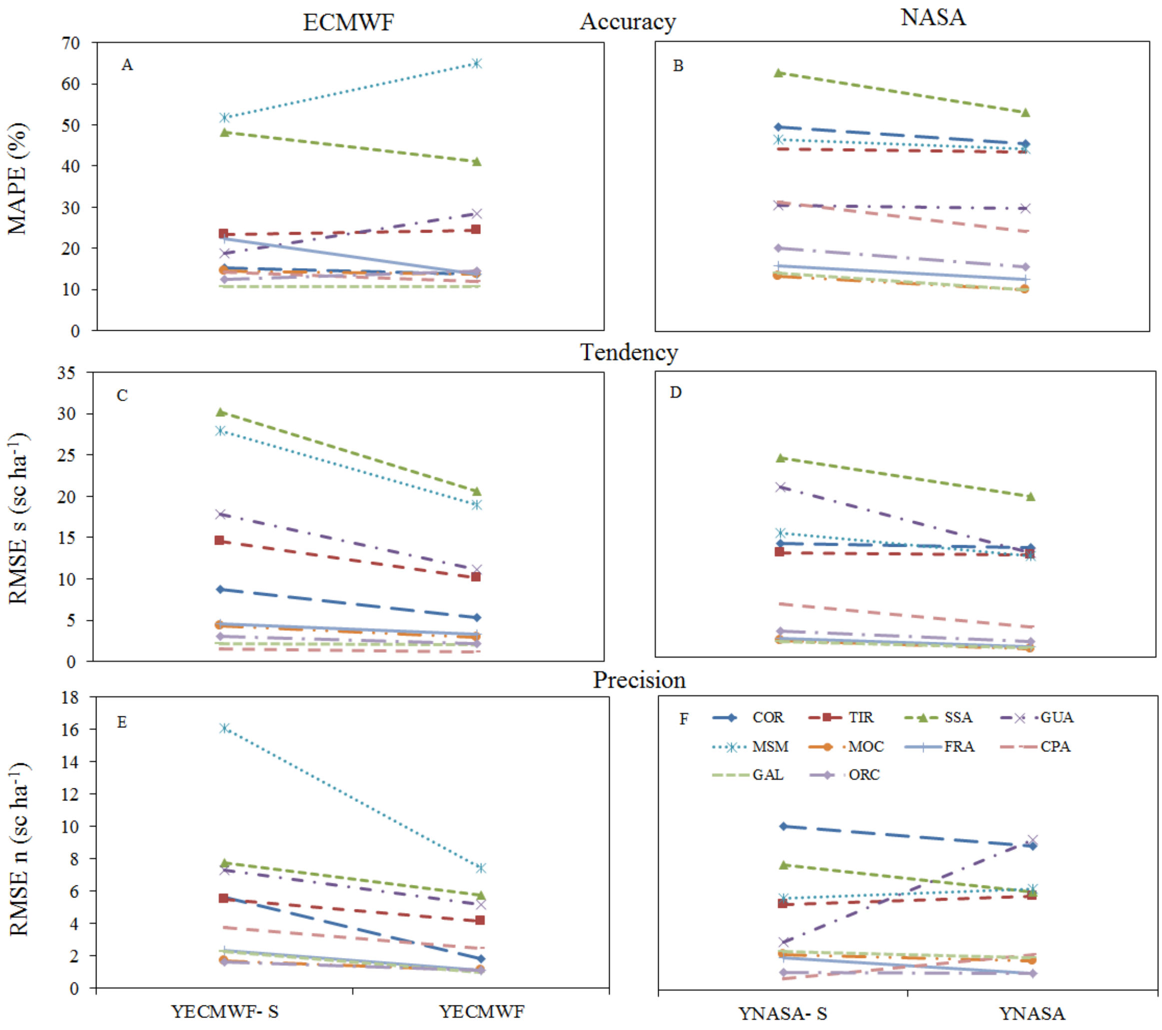

Figure 7. Evaluation of the accuracy, trend and accuracy of the ECMWF and NASA models calibrated with surface data (YECMWF-S, YNASA-S), and calibrated with model data (YECMWF and YNASA). Legend: COR- Coromandel; TIR- Tiros; SSA- Serra do Salitre; GUA- Guaxupé; MSM- Monte Santo de Minas; MOC- Mococa; FRA- Franca; CPA- Cristais Paulista; GAL-Gália; ORC- Osvaldo Cruz.

The penalizations for the factors evaluated were very similar between the models (Tables 7-9). The most penalized factors were water deficit and the influence of the previous year's yield, followed by minimum temperature (fTMIN). Maximum temperature (fTMAX) was penalized only in some locations such as Mococa and Gália for Y_{SC-S}; Coromandel, Franca and Cristais Paulista for Y_{ECMWF} and only Mococa for Y_{NASA}. The highest penalization was in Mococa at 8%, similar to that found by Santos and Camargo (2006), where high temperatures were penalized in Gália and Mococa, such as in 1994/1995 in Mococa with a yield loss of up to 6% due to an average temperature of 24.8 °C.

The highest penalization for fTMIN was 13% for Y_{SC-S} in Tiros, with absolute minimum temperatures of 0 °C in the 2002/2003, 2004/2005 and 2005/2006 harvests. Penalizations for fDH were highest (0.80) in Gália and Mococa, with a DEF of approximately 158 mm in the 2008/2009 harvest. This result is in agreement with Santos and Camargo (2006), in which Gália had the most intense and longest lasting DEFs, which caused larger yield losses.

Table 7. Percentage breakages of the models as a function of the penalties due to the yield of the previous year, water deficit, minimum temperature and maximum temperature, for Santos and Camargo model, parameterized with surface data (Y_{SC-S}).

Locations	Bieniality component: 1-[Ky0(Yaa/Yp)]	Hydric component: fDH	Thermal component fTMIN	fTMAX
		Superfície		
		Cerrado Mineiro		
Coromandel	0.53 ± 0.10	0.76 ± 0.052	0.30 ± 0.51	0
Tiros	0.73 ± 0.075	0.62 ± 0.07	12.06 ± 23.46	0
Serra do Salitre	0.42 ± 0.13	0.53 ± 0.031	0.36 ± 0.21	0
Média	0,56	0,63	4,24	0
		Sul de Minas		
Guaxupé	0.82 ± 0.064	0.80 ± 0.083	2.49 ± 5.81	0
Monte Santo de Minas	0.90 ± 0.057	0.49 ± 0.16	0,00	0
Mococa	0.48 ± 0.095	0.50 ± 0.094	0.95 ± 0.078	0.7 ± 0.9
Média	0,73	0,59	1,72	0,72
		Alta Mogiana		
Franca	0.85 ± 0.030	0.42 ± 0.056	0.65 ± 0.23	0
Cristais Paulista	0.38 ± 0.28	0.69 ± 0.084	2 ± 6.26	0
Média	0,61	0,55	1,30	0
		Noroeste Paulista		
Gália	0.16 ± 0.14	0.72 ± 0.10	13.89 ± 35.75	1.9 ± 10
Osvaldo Cruz	0.62 ± 0.036	0.47 ± 0.086	3 ± 6.26	0
Média	0,39	0,60	8,40	1,91

Legend: Average relative breaks ± Standard Deviation.

Table 8. Percentage model breaks due to the penalties due to previous year's yield, water deficit, minimum temperature and maximum temperature, for ECMWF (Y_{ECMWF}) model.

Locations	Byeniality component: 1-[Ky0(Yaa/Yp)]	Hydric component: fDH	Thermal component: fTMIN	fTMAX
		ECMWF		
		Cerrado Mineiro		
Coromandel	0.61 ± 0.09	0.48 ± 0.07	0.97 ± 0.030	1.74 ± 7.75
Tiros	0.69 ± 0.088	0.63 ± 0.06	0	0
Serra do Salitre	0.70 ± 0.13	0.51 ± 0.07	0	0
Média	0,66	0,54	0,97	1,74
		Sul de Minas		
Guaxupé	0.92 ± 0.024	0.48 ± 0.09	0	0
Monte Santo de Minas	0.92 ± 0.033	0.45 ± 0.11	0	0
Mococa	0.54 ± 0.08	0.45 ± 0.06	0.99 ± 0.002	0
Média	0,79	0,46	0,99	0
		Alta Mogiana		
Franca	0.55 ± 0.08	0.42 ± 0.07	0	0.76 ± 0.62
Cristais Paulista	0.80 ± 0.09	0.73 ± 0.05	0	0.82 ± 4.83
Média	0,67	0,57	0	0,79
		Noroeste Paulista		
Gália	0.20 ± 0.10	0.51 ± 0.11	0	0
Osvaldo Cruz	0.52 ± 0.049	0.52 ± 0.10	1.47 ± 1.72	0
Média	0,36	0,51	1,47	0

Legend: Average relative breaks ± Standard Deviation.

Table 9. Percentage model breaks as a function of the penalties due to the previous year's yield, water deficit, minimum temperature and maximum temperature, for the NASA (Y_{NASA}) model.

Locations	Bieniality component: 1-[Ky0(Yaa/Yp)]	Hydric component: fDH	Thermal component: fTMIN	fTMAX
		NASA		
		Cerrado Mineiro		
Coromandel	0.81 ± 0.038	0.46 ± 0.36	0.60 ± 0.29	0
Tiros	0.45 ± 0.15	0.74 ± 0.03	0.99 ± 0.003	0
Serra do Salitre	0.69 ± 0.13	0.54 ± 0.04	0.86 ± 0.18	0
Média	0,65	0,58	0,81	0
		Sul de Minas		
Guaxupé	0.75 ± 0.07	0.81 ± 0.06	0.81 ± 0.23	0
Monte Santo de Minas	0.78 ± 0.09	0.80 ± 0.05	0.58 ± 0.51	0
Mococa	0.38 ± 0.11	0.88 ± 0.03	0.87 ± 0.19	8 ± 25.2
Média	0,63	0,83	0,75	8,00
		Alta Mogiana		
Franca	0.79 ± 0.04	0.30 ± 0.09	0.73 ± 0.27	0
Cristais Paulista	0.88 ± 0.02	0.40 ± 0.045	0.84 ± 0.23	0
Média	0,83	0,35	0,78	0
		Noroeste Paulista		
Gália	0.30 ± 0.09	0.89 ± 0.03	0.57 ± 0.77	0
Osvaldo Cruz	0.51 ± 0.064	0.53 ± 0.085	0.99 ± 0.027	0
Média	0,4	0,71	0,78	0

Legend: Average relative breaks ± Standard Deviation.

The penalization factor for minimum temperature (fTMIN) in Cerrado Mineiro (Figure 8) was the highest for the determination of yield, second to penalization for maximum temperatures in the Y_{ECMWF} model. fTMIN and yield of the previous year (1-ky0 (Yaa/Yp)) were both penalized in Sul de Minas, except for the model with NASA data (Y_{NASA}), where the penalization for maximum temperature was higher in Mococa, as mentioned above. fTMIN was the most important factor in Alta Mogiana, except for Y_{ECMWF} that had no such penalization, which can be explained by the time scale (10 days) of the ECMWF data. In Noroeste Paulista, fTMIN had the highest percentages of breakage, followed by fTMIN.

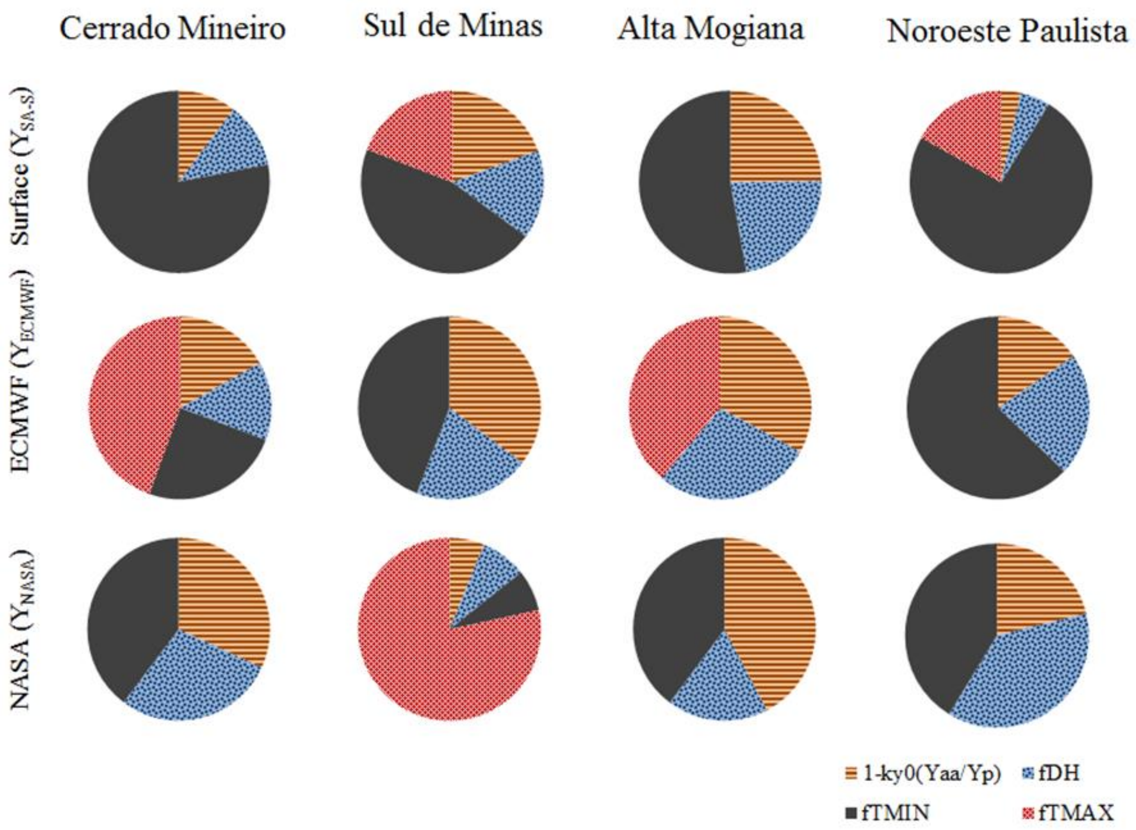

Figure 8. Percentage of relative breaks for the studied regions, according to different data sources (YSC-S), (YECMWF) and (YNASA). Caption: Yaa, penalization based on the previous year; FDH, penalization for water deficiency; FTMIN, penalization for minimum temperatures; FTMAX, penalization for maximum temperatures.

It was carried out the extrapolation of the accuracy data (MAPE) for the entire coffee region analyzed for a spatial understanding of Santos and Camargo model (2006) from the ECMWF and NASA grid data. In the west of the region (west of São Paulo) the accuracy of the models was higher with MAPE <28% for both Y_{ECMWF} (Figure 9.A) and Y_{NASA} (Figure 9.B). Cerrado Mineiro was the region with the lowest accuracy (MAPE> 44%). The places with the biggest problems were Monte Santo de Minas e Tiros, for both Y_{ECMWF} and Y_{NASA}.

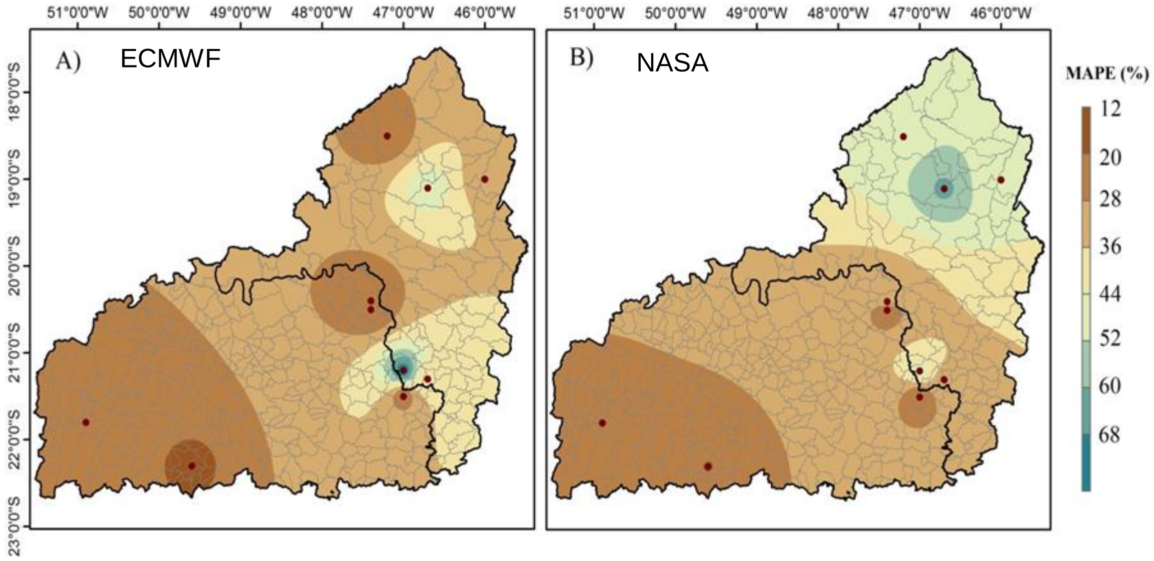

Figure 9. Interpolation of the MAPE values (%) in the region studied, of the ECMWF productivity models A) and NASA B).

The tests carried out for three new locations (Muzambinho, Rio Paranaíba and Passos) confirmed the improvement of the model when calibrated with the GD data. MAPE, RMSEs and RMSEn were lower for models calibrated with both ECMWF and NASA data than the model calibrated with surface data (Figure 10). Muzambinho responded the most to the calibration, with a MAPE of 46% when calibrated with surface data and 16% when calibrated with the NASA data. There was no such reduction only in Passos (MG), but the differences between the values calibrated with the surface and GD data were only 3 and 9% for ECMWF and NASA, respectively. The most accurate models were for Passos calibrated with ECMWF data and Muzambinho calibrated with NASA data, both with an RMSEn of 1.79 sacks ha^{-1}.

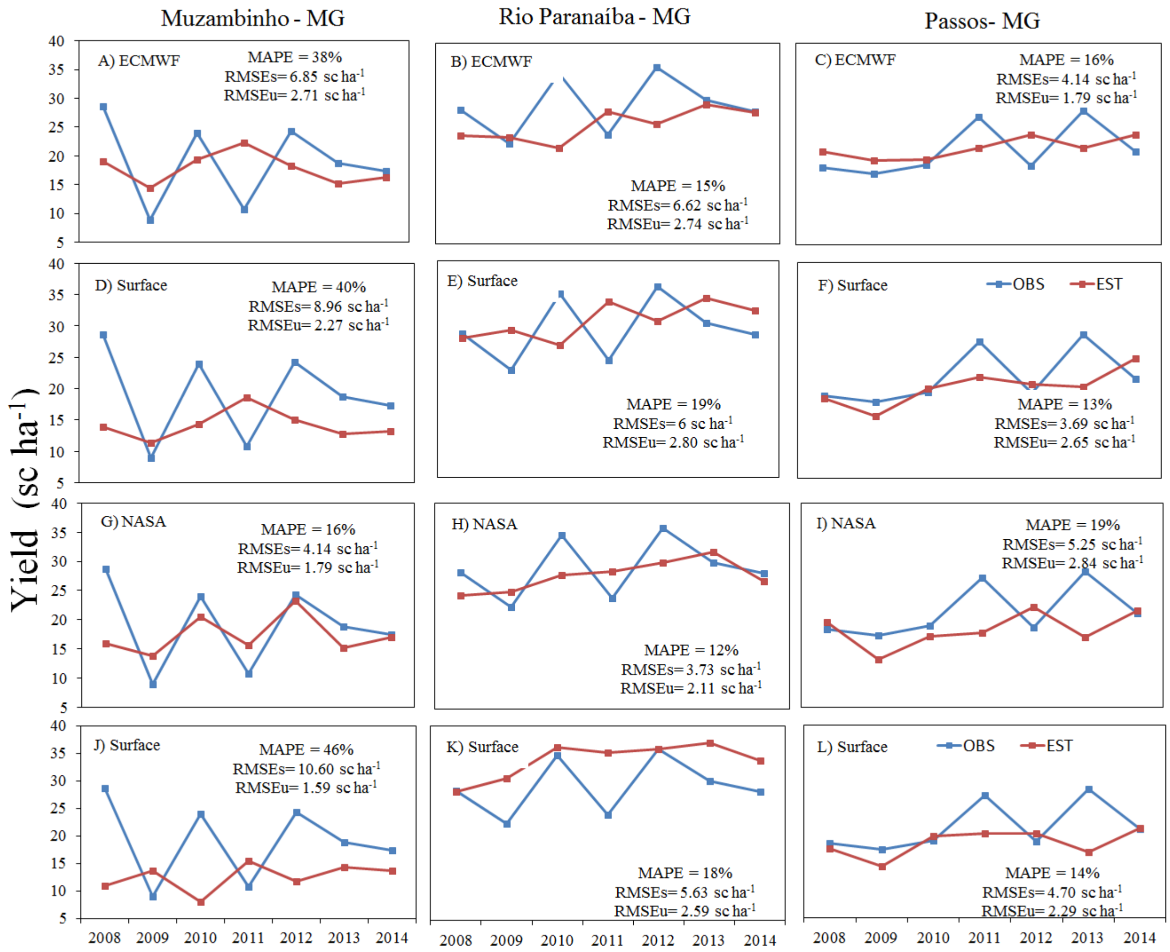

Figure 10. Comparison between observed (OBS) and estimated (EST) data from the tested locations for the two calibrated data sources; For Muzambinho: Y_{ECMWF}, $Y_{SC\text{-}S}$ (A, D); Y_{NASA}, $Y_{SC\text{-}S}$ (G, J); Rio Paranaíba: Y_{ECMWF}, $Y_{SC\text{-}S}$ (B, E); Y_{NASA}, $Y_{SC\text{-}S}$ (H, K); And Passos: Y_{ECMWF}, $Y_{SC\text{-}S}$ (C, F); Y_{NASA}, $Y_{SC\text{-}S}$ (I, L).

3.4 CONCLUSIONS

Coffee yield for São Paulo and Minas Gerais can be calibrated and estimated using the ECMWF and NASA grid data. The parametrization in the agrometeorological model, developed by Santos and Camargo (2006), provided accurate estimates of coffee yield in these two coffee-producing regions.

3.5 REFERENCES

AMARASINGHE, U. A.; HOANH, C. T.; D'HAEZE, D.; HUNG, T. Q. Toward sustainable coffee production in Vietnam: More coffee with less water. **Agricultural Systems,** v.136, p.96–105, 2015.

BAI, J.; CHEN, X.; DOBERMMAN, A.; YANG, H.; CASSMAN, K. G.; ZHANG, F. Evaluation of NASA satellite- and model-derived weather data for simulation of maize yield potential in China. **Agronomy Journal,** v.102, n.1, p. 9–16, 2010.

BECHTOLD. P.; KOHLER, M.; JUNG, T.; DOBLAS-REYES, F.; RODWELL, M. J.; VIART, F.; BALSAMO, G. Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decendal time-scales. **Q.J.R. Meteorol. Soc.** v.134, p.1337-1351, 2008.

Butt, M. S.; Sultan, M. T. Coffee and its consumption: benefits and risks. **Critical Reviews in Food Science and Nutrition**, Amherst, v.51, n.3, p.363-373, 2011.

CAMARGO, M. B. P. DE; SANTOS, M. A. DOS; BRUNINI, O.; FAHL, J. I.; MEIRELES, E. J. L.; LORENA, B. (2005) Teste de modelo agrometeorológico de monitoramento e de estimativa de produtividade do cafeeiro (*Coffea arabica* L.). no Estado de São Paulo. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 4. Londrina. **Anais...** Brasília, DF: Embrapa Café, 2005.

CAMARGO, M. B. P.; PEDRO JUNIOR, M. J.; ORTOLANI, A. A.; ALFONSI, R. R.; PINTO, H. S. Relações entre a precipitação pluviométrica e a produtividade do cafeeiro. **Ecossistema**, Espírito Santo do Pinhal, v. 9, v.1, p.166-171, 1984.

CAMARGO, M. B. P.; SANTOS, M. A.; PEDRO JUNIOR, M. J.; FAHL, J. I.; BRUNINI, O.; MEIRELES, E. J. L.; BARDIN, L. Modelo agrometeorológico de monitoramento e de estimativa de quebra de produtividade como subsidio à previsão de safra de café (*Coffea arabica* L.): resultados preliminares. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, Porto Seguro, 2003. **Anais...** Brasília, Embrapa Café 7576.

CAMARGO, A.P. **Balanço hídrico no Estado de São Paulo** 3 ed, Campinas, IAC, Bol. 116, 24p., 1971.

CARVALHO, L. G.; SEDIYAMA, G. C.; CECON, P. R.; RAMOS ALVES, H. M. Avaliação de um modelo agrometeorológico para previsão de produtividade de café em três localidades da região sul do Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, Santa Maria, v.11, n.2, p.343-352, 2003.

CHE, T.; DAI, L.; ZHENG, X.; LI, X.; ZHAO, K. Estimation of snow depth from passive microwave brightness temperature data in forest regions of northeast China. **Remote Sensing of Environment,** v.183, p.334-349, 2016.

CRAPARO, A. C. W.; VAN ASTEN, P. J. A.; LÄDERACH, P.; JASSOGNE, L. T. P.; GRABA, S. W. *Coffea arabica* yields decline in Tanzania due to climate change: Global implications. **Agricultural and Forest Meteorology** v.207, p.1–10, 2015.

CUBRY, P. et al. Global analysis of Coffea canephora Pierre ex Froehner (Rubiaceae) from the Guineo- Congolese region reveals impacts from climatic refuges and migration effects. **Genetic Resources Crop Evolution**, Witzenhausen v.60, p.483-501, 2013.

DAMATTA, F. M.; RAMALHO, J. D. C. Impacts of drought and temperature stress on coffee physiology and production: a review. **Brazilian Journal Plant Physiology**, Campos dos Goytacazes, v.18, p.55-81, 2006.

DAUZAT, J.; RAPIDEL, B.; BERGER, A. Simulation of leaf transpiration and sap flow in virtual plants: model description and application to a coffee plantation in Costa Rica. **Agric. Forest. Meteol**. v.109, p.143–160, 2001.

EVANGELISTA, A. W. P.; CARVALHO, L. G.; SEDIYAMA , G. C. Climatic zoning associated to the productive potential of coffee crop in Minas Gerais State, Brazil. **Revista Brasileira de Engenharia Agrícola Ambiental**, Campina Grande, v. 6, n. 3, p. 445-452, 2002.

GASPAR, N. A.; APARECIDO, L. E. O.; ROLIM, G. S.; SOUZA, P. S.; BOTELHOS, T. G. Software for the management of weather stations and for agrometeorological calculations. **Australian Journal of Crop Science**, v.9, n.6, 2015.

HOOGENBOOM, G. Contribution of agrometeorology to the simulation of crop production and its applications**. Agric. For. Meteor**., v.103, p.137–157, 2000.

KITZBERGER, C. S. G. et al. (2013) Diterpenes in green and roasted coffee of Coffea arabica cultivars growing in the same edapho-climatic conditions. **Journal of Food Composition and Analysis**, Campinas, v.30, p.52-57, 2013.

LIU, W. T. H.; LIU, B. W.Y. Comparação de três modelos de previsão de safra de café no Estado de Minas Gerais. **Ciência e Cultura**, São Paulo, v.40, n.8, p. 801–807, 1988.

MAHMOOD, R.; FOSTER, S. A.; LOGAN, D. The GeoProfile metadata, exposure of instruments, and measurement bias in climatic record revisited. **Int. J. Climatol.** v.26, p.1091–1124, 2006.

MARTINS, E.; APARECIDO, L. E. O.; SANTOS, L. P. S.; MENDONÇA, J. M. A.; SOUZA, P. S. (2015) Influencia das condições climáticas na produtividade e qualidade do cafeeiro produzido na região do sul de Minas Gerais. **Coffe Science**, Lavras, v.10, n.4, p.499-506.

MOELETSI, M. E; WALKER, S. Evaluation of NASA Satellite and Modelled Temperature Data for Simulating Maize Water Requirement Satisfaction Index in the Free State Province of South Africa. **Physics and Chemistry of the Earth,** Parts A/B/C, p.50-52 p.157-164, 2012. Disponível em: http://dx.doi.org/10.1016/j.pce.2012.08.012.

MONTOYA, R. E.; ARCILA, J.; JARAMILLO, A.; RIA˜NO, N.; QUIROGA, F. **Modelo para simular la producción potencial del cultivo de café en Colombia.** Centro Nacional de Investigaciones del café “Pedro Uribe Mejía”, Chinchiná, Colombia, 2009.

PEREIRA, A. R.; ANGELOCCI, L. R.; SENTELHAS, P. C. Agrometeorologia: fundamentos e aplicações práticas. Guaíba: Agropecuária, 478p, 2002.

PERSSON, A. User guide to ECMWF forecast products**. Meteorological Bull**. M3.2, ECMWF, Reading, United Kingdom, p.115, 2001.

PICINI, A. G.; CAMARGO, M. B. P.; ORTOLANI, A. A.; FAZUOLI, L. C.; GALLO, P. B. Desenvolvimento e teste de modelos agrometeorológicos para a estimativa de produtividade do cafeeiro. **Bragantia**, Campinas, v.58, n.1, p.157-170, 1999.

REVISTABW 2015. Probabilidade e Estatística: Acurácia, Precisão e Exatidão. **Revista Brasileira de Web: Tecnologia**. Disponível em http://www.revistabw.com.br/revistabw/probabilidade-e-estatistica-acuracia-precisao-e-exatidao/. Visitado em: 11/12/2016.

RODRÍGUEZ, D. et al. A coffee agroecosystem model: II., dynamics of coffee berry borer. **Ecological Modelling**, Towson, v.248, p.203-214, 2013.

RODRÍGUEZ. D.; CURE, J. R.; COTES, J. M.; GUTIERREZ, A.P.; CANTOR, F. A coffee agroecosystem model: I. growth abd development of the coffee plant. **Ecological Modelling** v.222, p.3626-3639, 2011.

ROLIM, G. S.; RIBEIRO, R. V.; AZEVEDO, F. A.; CAMARGO, M. B. P.; MACHADO, E. C. Predicion of fruits number from the quantity of reproductive structures of orange. **Rev Bras Frutic.** v.30, n.1, p.48-53, 2008.

SANTOS, M. A.; CAMARGO, M. B. P. Calibration of an agrometeorological model for predicting coffee (*Coffea arabica* L.) productivity in Sao Paulo state, Brazil. **Bragantia**, v.65, n.1, p.173-183, 2006.

SANTOS, M.A. dos. **Parametrização de coeficientes de sensibilidade e teste de modelos agrometeorológicos de estimativa de quebra de produtividade do cafeeiro (*Coffea arabica* L.).** Campinas, 146p. Dissertação (Mestrado), Instituto Agronômico de Campinas, APTA, 2005.

SILVA, G. L. S. P.; VICENTE, J. R.; CASER, D. V. Avaliação de previsões de rendimento de culturas no Estado de São Paulo, fornecidas por modelos agrometeorológicos, anos agrícolas 1984/85 e 1985/86. **Informações Econômicas**, São Paulo, v.17, n.6, p.61-45, 1987.

THORNTHWAITE, C. W.; MATHER, J. R. The water balance. Centerton: Drexel Institute of Technology, **Climatology**. V.8, n.1, p.104, 1955.

TOSELLO, R. N.; ARRUDA, H. V. de. Correlação entre estimativas oficiais de produção de café e precipitação pluviométrica, no Estado de São Paulo. **Bragantia**, Campinas, v.21, p.449-465, 1962.

VASCONCELLOS, E. B. de C.; ROMANI, L. A. S. (2012) Validação do modelo de estimativa de produtividade do cafeeiro (Arábica) na região sul do Estado de Minas Gerais. In: Mostra de estagiários e bolsistas da Embrapa Informàtica Agropecuaria, 8. Campinas. **Resumos...** Brasília, DF: Embrapa, p.139-142, 2012.

WANG, N.; JASSOGNE, L.; VAN ASTEN, P. J. A.; MUKASA, D.; WANYAMA, I.; KAGEZI, G.; GILLER, K. E. Evaluating coffee yield gaps and important biotic, abiotic, and management factors limiting coffee production in Uganda. **European Journal of Agronomy**, v.63, p.1-11, 2015. Disponível em: http://dx.doi.org/10.1016/j.eja.2014.11.003

WEILL, M. A. M. **Avaliação de fatores edafoclimáticos e do manejo na produção de cafeeiros (*Coffea arabica* L.) na região de Marília e Garça, SP**. 182 f. Dissertação (Mestrado). Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 1990.

WHITE, J. F.; HOOGENBOOM, G.; STACKHOUSE JR, P. W.; HOELL, J.M. Evaluation of NASA satellite- and assimilation model-derived long-term daily temperature data over the continental US. **Agricultural and Forest Meteorology** v.148, p.1574– 1584, 2008.

WU, H.; HUBBARD, K. G.; YOU, J. Some concerns when using data from the Cooperative Weather Station Networks: a Nebraska case study. **J. Atmos. Ocean. Technol.** v.22, p.592–602, 2005.