

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE ASTRONOMIA, GEOFÍSICA E CIÊNCIAS ATMOSFÉRICAS  
DEPARTAMENTO DE CIÊNCIAS ATMOSFÉRICAS

RODRIGO LUSTOSA

**Estimativa da temperatura do ar com sensoriamento remoto e estações meteorológicas na  
Região Metropolitana de São Paulo**

SÃO PAULO

2021

RODRIGO LUSTOSA

**Estimativa da temperatura do ar com sensoriamento remoto e estações meteorológicas na  
Região Metropolitana de São Paulo**

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Meteorologia da Universidade de São Paulo, como requisito necessário à obtenção do título de Bacharel em Meteorologia.

Orientador: Prof. Dr. Humberto Ribeiro da Rocha

SÃO PAULO

2021

## RESUMO

A temperatura do ar próxima da superfície ( $T_a$ ) é uma das variáveis mais importantes da meteorologia e climatologia, mas atualmente não é medida por sensoriamento remoto, que possui uma ampla cobertura da superfície, e sim apenas por estações meteorológicas, pontos isolados no globo. Considerando que a atmosfera próxima da superfície possui expressivas trocas de calor com a superfície, esperam-se boas correlações entre  $T_a$  e variáveis de superfície. Mas como a atmosfera é fluida, uma parcela de ar não será influenciada apenas pela superfície sobre a qual se encontra mas também por todas sobre as quais passou, com um peso maior para as mais recentes.

Este estudo visou analisar as relações empíricas entre  $T_a$  e a temperatura de superfície ( $T_S$ ), obtida dos satélites Landsat 5 e 8 (horário de revisita aproximadamente às 10 h no horário local). Especificamente, procurou-se compreender como o entorno influencia a correlação entre  $T_S$  e  $T_a$  em estudo de caso na Região Metropolitana de São Paulo (RMSP), utilizando dados de 2008 até 2021. O efeito do entorno foi quantificado calculando-se a média de  $T_S$  dentro de diferentes raios ao redor de estações meteorológicas. Para cada raio e cada estação, calculou-se a correlação entre  $T_S$  e  $T_a$ , com distinção sazonal. A distinção sazonal foi feita separando os dados em fotoperíodo ( $N$ ) maior e menor que 12 h, que representam a primavera e verão e o outono e inverno astronômicos, respectivamente.

Constata-se que  $T_S$  e  $T_a$  se relacionam de forma linear, que a correlação é maior para  $N < 12$  h e que a variação espacial da correlação é maior para  $N > 12$  h. Ao considerar o entorno, em nem todas as estações meteorológicas é possível melhorar a correlação, mas a melhora é significativa nas demais. O raio das circunferências que maximizam a correlação muda de estação para estação, e suas escalas passam tanto pela micro quanto pela mesoescala da meteorologia. Muitas estações possuem correlações com comportamentos similares em função do raio, e é possível que isso esteja relacionado com as características do entorno das estações. Uma análise preliminar dessa hipótese foi feita para a estação do IAG Cientec, localizada no Parque do Estado, e constatou-se que a circunferência de máxima correlação, de raio igual a 750 m, praticamente tangencia a fronteira entre o parque e a área urbana ao redor, ou seja, a correlação aumenta com o raio até 750 m e a partir daí os valores de  $T_S$  da área urbana passam a ser considerados, ao mesmo tempo em que a correlação diminui.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>4</b>
<b>2</b>	<b>OBJETIVO</b>	<b>8</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>8</b>
3.1	ÁREA DE ESTUDO	8
3.2	PERÍODO DE ESTUDO	8
3.3	CONJUNTO DE DADOS	8
3.3.1	Estações meteorológicas	8
3.3.2	Satélites Landsat	9
3.4	MODELOS	11
3.4.1	Técnicas Estatísticas	11
3.4.2	Fotoperíodo	11
3.5	MÉTODOS	12
3.5.1	Área de influência de $T_S$ sobre $T_a$	13
<b>4</b>	<b>RESULTADOS E ANÁLISES</b>	<b>13</b>
4.1	Distribuição entre $T_a$ e $T_S$	13
4.2	Influência do Entorno na Correlação	16
4.3	Comparação Preliminar Entre Composição do Entorno e Correlação	22
<b>5</b>	<b>CONCLUSÕES E RECOMENDAÇÕES</b>	<b>23</b>
	<b>REFERÊNCIAS</b>	<b>25</b>
	<b>APÊNDICE A – Disponibilidade de dados de <math>T_a</math> por estação</b>	<b>27</b>
	<b>APÊNDICE B – Dispersão entre <math>T_a</math> e <math>T_S</math> para cada estação meteorológica</b>	<b>28</b>
	<b>APÊNDICE C – Dispersão entre <math>T_a</math> e <math>T_S</math> para cada estação meteorológica</b>	<b>32</b>

## 1 INTRODUÇÃO

A temperatura do ar próxima da superfície ( $T_a$ ), geralmente medida a 2 m do solo, está entre as variáveis mais importantes da meteorologia e climatologia, com utilidade indispensável na previsão do tempo, agronomia, estudos de mudanças climáticas, entre outros. Suas aplicações atingem até áreas interdisciplinares como saúde e planejamento urbano.

Embora elementar, mesmo com o desenvolvimento tecnológico e científico das últimas décadas, em especial o crescente uso de dados de satélites artificiais,  $T_a$  ainda é medida majoritariamente por estações meteorológicas, que são pontos isolados no globo e sua abrangência fica ainda mais comprometida em locais pouco acessíveis, como a floresta tropical Amazônica, ou países economicamente vulneráveis, com poucos recursos para esse tipo de investimento.

Enquanto isso, dados de sensoriamento remoto (i.e., de sensores espectrais a bordo de satélites, estações espaciais e afins) já provêm informações importantes do planeta, como dados de cobertura vegetal, precipitação, vento e temperatura da superfície ( $T_S$ ). Em contraste com  $T_a$ , dados de sensoriamento remoto possuem ampla cobertura da superfície, independente de acessibilidade ou economia local, e se limitam a sua órbita, período e horário de revisita (para satélites não-geoestacionários) e resolução espacial e espectral. Por exemplo, no satélite Landsat 8, os sensores Imageador Operacional de terra (OLI — *Operational Land Imager*) e Sensor de Infravermelho Termal (TIRS — *Thermal Infrared Sensor*) possuem respectivamente resolução espacial de 30 e 100 m (que é reamostrado para 30 m) somando 11 bandas espectrais, possui revisita de 16 dias e passa em horário fixo, aproximadamente às 10 horas da manhã no horário solar local (LANDSAT..., ).

A obtenção de  $T_a$  por sensoriamento remoto com boa precisão é um processo complicado, pois em contraste com nuvens, chuva e superfície, o ar possui uma baixa interação com a radiação eletromagnética, único elemento que sensores remotos medem diretamente. Embora ocorra uma boa interação no infravermelho termal (TIR — *Thermal Infrared*), em especial pelo vapor d'água, as camadas superiores da atmosfera mascaram as inferiores e as características do ar próximo da superfície se perdem (STULL, 2015). Entretanto, a atmosfera recebe energia principalmente da radiação solar e da superfície (PETTY, 2006), onde esta última é quem mais contribui com energia para o ar próximo da superfície, por calor sensível e latente, e dessa forma  $T_a$  estará inter-relacionada com  $T_S$  e outras variáveis de superfície, como já observado por Sheng et al. (2017), Good et al. (2017), Yan et al. (2014), Yokobori e Ohta (2009), Cheng et al. (2008).

Dentre as análises da relação entre  $T_S$  e  $T_a$ , Cheng et al. (2008) demonstraram teórica e empiricamente que essa relação é dependente do tipo de cobertura da superfície. Fizeram um

estudo piloto sobre os efeitos de mudança do uso da terra em  $T_a$  no norte do Taiwan nos dias 16 de março e 4 de abril de 2005 (primavera) para quatro tipos de cobertura: corpos d'água, áreas construídas, plantações de arroz e outros tipos de vegetação. A classificação foi feita com imagens SPOT-4 sensor Visível e Infravermelho de Alta Resolução (HRVIR — *High Resolution Visible Infrared*), 20 m de resolução, e mediram *in situ*  $T_S$  e  $T_a$  em seis alturas (0,2, 0,4, 0,8, 1,2 e 2,0 m) em diferentes locais na área de estudo com um termômetro infra-vermelho entre às 11:00 e às 14:30 no horário local. Dos dados *in situ*, mostraram que para cada tipo de cobertura a relação entre  $T_a$  em 2,0 m e  $T_S$  é linear e observaram que  $T_S$  é no geral maior que  $T_a$ , exceto para corpos d'água, onde o inverso ocorre. Para áreas de 1,1 km por 1,1 km, estudaram como diferentes tipos de mudança de uso da terra modificariam  $T_a$ , admitindo uma  $T_S$  e  $T_a$  característica para cada tipo de cobertura de superfície. Observou-se que aumentar a cobertura de áreas construídas aumenta categoricamente  $T_a$  média, independente de como se modificam os outros tipos de cobertura.

Good et al. (2017) estudaram as relações entre  $T_S$  e  $T_a$  mensais utilizando dados globais de 17 anos.  $T_S$  foi obtida dos sensores Radiômetro de Varredura ao Longo do Percorso (ATSR — *Along-Track Scanning Radiometer*), com passagem às 10:00/22:00 e às 10:30/22:30 no horário solar local e resolução espacial de  $0,05^\circ$  ( $\simeq 5$  km), e  $T_a$  são valores interpolados de estações meteorológicas, disponíveis em grade no conjunto de dados Série Temporal da Unidade de Pesquisa Climática (CRU TS — *Climate Research Unit Time Series*) com a mínima mensal e a máxima mensal. O foco foi o comportamento da diferença entre as duas temperaturas ( $T_S$  diurna menos  $T_a$  máxima e  $T_S$  noturna menos  $T_a$  mínima), de acordo com trimestre, localização, vegetação, uso de terra e altitude. Constatou-se que os padrões das duas variáveis são muito similares e com correlação ( $r$ , Coeficiente de Pearson) próxima de 1 (exceto na região equatorial), e  $T_S$  é no geral maior que  $T_a$ , com exceções nos polos e florestas tropicais. A maior diferença foi entre  $T_S$  diurna e  $T_a$  máxima. A variação da diferença com a porcentagem de cobertura vegetal é sempre negativa e é mais intensa em regiões equatoriais e nos dados diurnos. A diferença em valores diurnos também é mais influenciada por diferentes tipos de uso de terra que os noturnos. Por fim, a correlação entre as duas variáveis diminui com o aumento da altitude. Com esses resultados, os autores ainda argumentam sobre e estimulam o uso de  $T_S$  para estudos climáticos e de tempo, o que ainda é pouco praticado, ao contrário da  $T_a$ , amplamente utilizada.

Sheng et al. (2017) quantificaram relações temporais e espaciais entre  $T_a$  e  $T_S$  em diferentes tipos de cobertura de superfície em estudo de caso no norte da província Zeijiang numa área de 105 400 km<sup>2</sup>, costa sudeste da China, que possui um clima de monções e com altitude variando entre  $-6$  e 1170 m, e no período de um ano, entre 1 de janeiro de 2014 e 31 de dezembro de 2014. Para isso, utilizaram dados de  $T_a$  de 601 estações meteorológicas automáticas

do Escritório Meteorológico Zhejiang e de  $T_S$  do satélite Terra sensor Radiômetro Espectral de Imagem de Resolução Moderada (MODIS — *Moderate-Resolution Imaging Spectroradiometer*), horário solar local às 10:30 e 22:30, produto *Terra MODIS Version 5 daily level 3 TS* (MOD11A1), 0,928 km de resolução espacial, que usa o algoritmo *Generalized Split-Window* e totalizou 283 imagens diurnas e 237 noturnas (não atingiu 365 por contaminação de nuvens). Também utilizaram dados Landsat 8 sensor OLI, resolução 30 m, de 2013 e 2014, para determinar a cobertura da área de estudo dividindo em três grupos, áreas de superfície impermeável (ASI), vegetação e corpos d'água. Constataram que há sazonalidade coincidente em  $T_S$  e  $T_a$  para todos os tipos de cobertura e para o dia e noite. Para a média das duas variáveis nessas categorias,  $T_S$  é majoritariamente menor que  $T_a$  de mesmo horário, com exceções ocorrendo para ASI durante o dia em torno do inverno para as outras duas coberturas (dia e noite). Para quantificar quão possível se é prever  $T_a$ , eles realizaram regressões lineares de  $T_a$  por  $T_S$  usando 70 % dos dados disponíveis separando por horário, cobertura e estação do ano (trimestres) e compararam com os 30 % restantes. Constatou-se que a inclinação dos ajustes é majoritariamente menor que 1 ( $T_a$  cresce mais lentamente que  $T_S$ ). A correlação entre  $T_a$  estimada e observada para os 30 % restantes, sempre positiva, no geral diminui no verão e diminui durante o dia (sobre corpos d'água também diminui no inverno). Para a vegetação e ASI, durante o dia  $r$  fica entre 0,808 e 0,887 (excluindo-se o verão, onde é igual a 0,122 e 0,472) e durante a noite fica entre 0,851 e 0,948 (excluindo-se novamente o verão em ASI, igual a 0,778). Para corpos d'água a correlação é majoritariamente menor, com valor máximo de 0,747 durante o dia e 0,824 durante a noite. A Raiz do Erro Quadrático Médio (RMSE — *Root Mean Square Error*) entre  $T_a$  estimada e observada é no geral menor durante a noite (de 1,53 a 2,66 K) do que durante o dia (de 2,06 a 4,31 K). Por fim, considerando que o ar é influenciado não só pelo pixel onde se encontra a estação mas também pelo seu entorno, os autores analisaram o efeito de se realizar médias de  $T_S$  nos pixels em torno da estação na correlação final (utilizaram quadrados centrados em cada estação, com comprimento variando de 1 a 75 pixels, aproximadamente de 1 a 75 km) e observaram que a correlação aumenta com o tamanho do quadro para ASI e para corpos d'água (vegetação fica praticamente constante). Esse aumento fica cada vez menor com o tamanho do quadro e com assíntota aproximadamente entre 0,96 e 0,99 (de acordo com horário e tipo de cobertura). Não foi encontrado um ponto onde a partir dele a correlação diminuía.

A Região Metropolitana de São Paulo (RMSP), importante área econômica brasileira, se encontra no Sudeste do Brasil e, como o sudeste da China, possui clima de monções (REBOITA et al., 2010). Ferreira (2019) analisou os padrões de  $T_S$  e vegetação na RMSP com dados do satélite Aqua sensor MODIS (250 m a 1 km e às 10:00/22:00) e de morfologia urbana da classificação Zonas Climáticas Locais (LCZ — *Local Climate Zones*), de 2002 a 2017, e com distinção

entre estação chuvosa e seca. Com as LCZ, Ferreira constatou que a maior parte da área urbana da RMSP é composta por construções baixas (casas) e adensadas. Observou também que a área urbana é sempre mais quente que os arredores (fenômeno de ilha de calor), embora no período diurno alguns bairros centrais específicos possuem temperatura significativamente menor que a área urbana ao redor. Segundo a classificação LCZ, esses bairros possuem estruturas ou altas e adensadas (prédios) ou baixos/médios e espaçados. Há também um maior Índice de Vegetação da Diferença Normalizada (NDVI) em comparação com outros bairros. Esse comportamento não ocorre durante a noite, que possui apenas um centro mais quente e que esfria em direção à fronteira urbana. Os padrões espaciais da estação chuvosa e seca são similares para um mesmo horário, apenas com a estação seca possuindo temperaturas mais frias. A  $T_S$  diurna possui uma boa correlação negativa com o NDVI, com uma correlação maior para a resolução de 1 km ( $R^2 = 0,90$ , onde  $R^2$  é o Coeficiente de Determinação) do que para de 250 m ( $R^2 = 0,88$ ). Valores noturnos apresentaram correlações menores.

Existe atualmente uma plataforma online para análise de dados de sensoriamento remoto e similares em nuvem e com um amplo banco de dados conhecido como *Google Earth Engine* ou GEE (GORELICK et al., 2017). Nele há dados dos satélites Landsat, que obtêm dados espectrais da superfície com 30 m de resolução há mais de três décadas, porém o GEE ainda não possuía um produto de  $T_S$  para eles quando este trabalho foi feito. Dessa forma, Ermida et al. (2020) produziram um código aberto para GEE que computa  $T_S$  utilizando dados Landsat em conjunto com dados de reanálise de coluna de vapor de água e dados de emissividade de superfície, também disponíveis no GEE. Foi utilizado um algoritmo empírico calibrado com dados de simulações e checado com dados de estações *in situ*.

Compreender a fundo as relações empíricas entre  $T_a$  e variáveis de superfície pode permitir realizar estimativas de  $T_a$  por sensoriamento remoto e conhecer seu grau de precisão. Outra vantagem seria sequer precisar utilizar  $T_a$  em estudos onde ela é normalmente utilizada, bastando utilizar outras variáveis (e.g.,  $T_S$ ), desde que a relação desta com  $T_a$  seja conhecida. Por exemplo, em análises de ilha de calor urbano já existem métodos consistentes onde se utiliza  $T_S$  no lugar de  $T_a$ , e.g. Li et al. (2018), que sugerem uma regressão linear entre  $T_S$  e a fração de área de superfície impermeável (ou seja, fração de área urbanizada) e a intensidade da ilha de calor será o coeficiente angular do ajuste. Sua consistência permite comparar resultados de diferentes áreas urbanas, o que não é possível com dados de  $T_a$ , já que o grau de distribuição de estações numa área urbana e arredores pode influenciar no resultado. Utilizar variáveis diretamente obtidas por sensoriamento remoto e conhecer as relações com  $T_a$  permite ao mesmo tempo comparar resultados de diferentes locais e estimar o impacto em  $T_a$ .



## 2 OBJETIVO

Este trabalho visa contribuir no entendimento das relações empíricas entre  $T_a$  e  $T_S$  em um estudo de caso na RMSP. Para isso, utilizaram-se dados de alta resolução dos satélites Landsat 5 e 8 para se obter variáveis de superfície e dados de estações meteorológicas automáticas para variáveis atmosféricas. Não se encontrou na literatura estudos sobre a relação entre  $T_a$  e  $T_S$  horários em um período maior que um ano. Aqui pretendeu-se abranger períodos maiores de forma a incluir possíveis diferenças entre anos, se aproveitando da longa série temporal dos satélites Landsat e das estações meteorológicas na área de estudo.

Especificamente, o objetivo foi verificar possíveis impactos do entorno na correlação entre  $T_a$  e  $T_S$ , distinguindo a sazonalidade. O impacto do entorno é quantificado calculando-se a média de  $T_S$  em diferentes raios ao redor das estações meteorológicas e para cada estação e raio calcula-se a correlação com  $T_a$ .

## 3 METODOLOGIA

### 3.1 ÁREA DE ESTUDO

A Região Metropolitana de São Paulo (RMSP) está localizada no Sudeste do Brasil, região de clima de monções (REBOITA et al., 2010), possui a maior conurbação da América do Sul e a 5ª maior do mundo em população, com aproximadamente 22,5 milhões de habitantes (DEMOGRAPHIA, 2021). Abrange uma área em torno de 7965 km<sup>2</sup> e sua principal cobertura vegetal é a Mata Atlântica.

### 3.2 PERÍODO DE ESTUDO

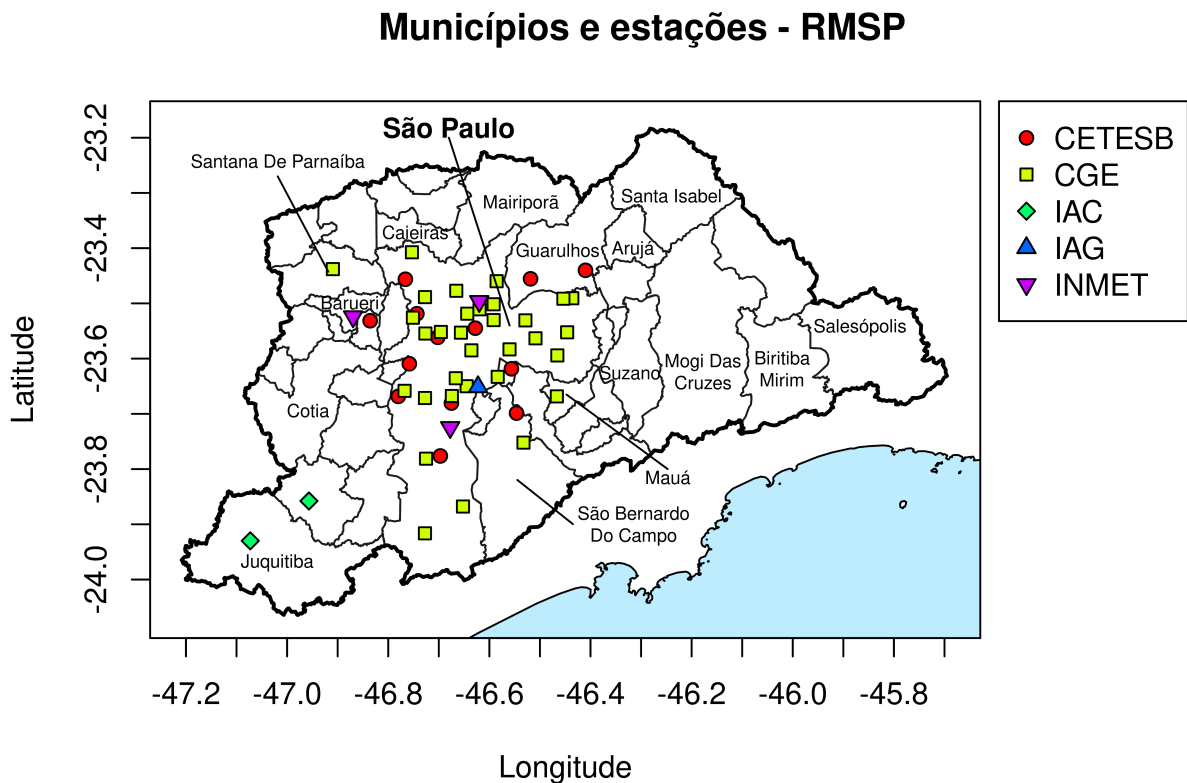
Utilizaram-se dados a partir de 01 de janeiro de 2008 e até 30 de junho de 2021.

### 3.3 CONJUNTO DE DADOS

#### 3.3.1 Estações meteorológicas

A variável  $T_a$  é obtida de 51 estações meteorológicas automáticas espalhadas pela RMSP (Figura 1). Elas pertencem à cinco gestores diferentes, que são o Centro de Gerenciamento de Emergências Climáticas da Prefeitura de São Paulo (CGE), 32 estações, a Companhia Ambiental do Estado de São Paulo (CETESB), 13 estações, o Instituto Nacional de Meteorologia (INMET), 3 estações, o Centro Integrado de Informações Agrometeorológicas do Instituto Agrônomo (CIIAGRO-IAC), 2 estações, e o Instituto de Astronomia, Geofísica e Ciências Atmosféricas da

Universidade de São Paulo (IAG-USP), 1 estação. Como mostra a Figura 1, sua distribuição não é homogênea e se concentra na área urbana. A disponibilidade de dados no tempo também não é uniforme, com algumas estações possuindo mais dados que outras e a maior disponibilidade ocorre a partir de 2012, quando dezenas dessas estações entraram em operação (Figura A1).



**Figura 1:** Divisão municipal da Região Metropolitana de São Paulo (RMSP) e estações meteorológicas utilizadas, classificadas por gestor.

### 3.3.2 Satélites Landsat

Os satélites Landsat 5 sensor Mapeador Temático (TM — *Thematic Mapper*) e Landsat 8 sensor OLI/TIRS possuem resolução espacial de 30 m para a onda curta, do visível (VIS) ao infravermelho de onda curta (SWIR — *Short-Wave Infrared*), e respectivamente 120 e 100 m para a onda longa, no TIR, que são reamostrados para 30 m (LANDSAT..., ). Passam em torno das 10:00 no horário solar local e tem período de revisita de 16 dias. Seus dados são consistentes entre si. Suas bandas são utilizadas para estimar  $T_S$ . Para as bandas de radiação de onda curta, se utiliza a refletância de superfície, que é uma refletância com correção atmosférica e disponível no GEE. Para a radiação de onda longa, utiliza-se sua temperatura de brilho no topo da atmosfera ( $T_b$ ).

O NDVI, índice espectral de vegetação, é utilizado no algoritmo para computar  $T_S$  e é dado pela equação (1), onde  $r_{espectro}$  é a refletância do pixel no descrito espectro de radiação (SZABÓ; GÁCSI; BALÁZS, 2016). O NDVI considera que quanto mais saudável a vegetação está e quanto mais vegetação há, maior a refletância no NIR em relação ao vermelho. Ele adquire valores de  $-1$  a  $1$ , onde quanto mais próximo de  $1$  maior a presença da vegetação. Valores positivos pequenos representam área urbana, solo exposto e podem representar corpos d'água e valores negativos no geral representam água.

$$\text{NDVI} = \frac{r_{NIR} - r_{vermelho}}{r_{NIR} + r_{vermelho}} \quad (1)$$

$T_S$  é computada no GEE segundo Ermida et al. (2020), que utiliza o algoritmo empírico *Statistical Mono-Window* (SMW), a partir da equação (2).

$$T_S = A_i \frac{T_b}{\varepsilon} + B_i \frac{1}{\varepsilon} + C_i \quad (2)$$

onde  $\varepsilon$  é a emissividade no pixel e  $A_i$ ,  $B_i$  e  $C_i$  são parâmetros obtidos empiricamente por Ermida et al. (2020) e estão em função da coluna total de vapor d'água (TCWV). TCWV vem de dados de reanálise do *National Centers for Environmental Prediction/National Center for Atmospheric Research* (NCEP/NCAR) e  $\varepsilon$  é estimada a partir do NDVI obtido pelo Landsat e de dados de emissividade média ( $\varepsilon_{\text{médio}}$ ) e NDVI médio ( $\text{NDVI}_{\text{médio}}$ ) do Banco de Dados de Emissividade Global ASTER (ASTER-GED), onde ambos vem de imagens do satélite Terra sensor Radiômetro Avançado de Emissão Térmica e Reflexão a Bordo de Espaço (ASTER) de céu limpo, resolução de 100 m, de 2000 até 2008. Essa estimativa utiliza as equações (3) e (4). Primeiro calcula-se (3) usando  $\text{NDVI}_{\text{médio}}$ , substitui-se o resultado em (4) e calcula-se  $\varepsilon_{\text{veg}}$  característico do pixel usando  $\varepsilon_{\text{médio}}$  como  $\varepsilon$ . Em seguida recalcula-se (3) e (4), dessa vez utilizando o NDVI obtido pelo Landsat e o  $\varepsilon_{\text{veg}}$  há pouco obtido.

$$\text{FVC} = \left( \frac{\text{NDVI} - \text{NDVI}_{\text{solo nu}}}{\text{NDVI}_{\text{veg}} + \text{NDVI}_{\text{solo nu}}} \right)^2 \quad (3)$$

onde FVC é a fração de cobertura vegetal (teoricamente de 0 a 1),  $\text{NDVI}_{\text{solo nu}} \equiv 0,2$  é o NDVI típico de solo nu e  $\text{NDVI}_{\text{veg}} \equiv 0,86$  é o típico para vegetação.

$$\varepsilon = \text{FVC} \cdot \varepsilon_{\text{veg}} + (1 - \text{FVC}) \varepsilon_{\text{solo nu}} \quad (4)$$

onde  $\varepsilon_{\text{solo nu}}$  é a emissividade do pixel para o caso  $\text{NDVI} = \text{NDVI}_{\text{solo nu}}$  e  $\varepsilon_{\text{veg}} \equiv 0,99$  é para o caso  $\text{NDVI} = \text{NDVI}_{\text{veg}}$ .

A calibração de  $A_i$ ,  $B_i$  e  $C_i$  foi feita com um banco de dados global de simulações de temperatura de brilho ( $T_b$ ), perfil atmosférico, emissividade e  $T_S$  e considerando as propriedades dos sensores de cada satélite. Com os valores de simulações, observa-se que o erro das estimativas é pequeno para valores baixos de TCWV e aumenta com seu aumento, onde o percentil 95 do erro vai de aproximadamente 1 K em TCWV  $\simeq$  18 cm para 5 K em TCWV  $\simeq$  60 cm. O viés do Landsat 5 também aumenta com o aumento de TCWV, mas é satisfatório para valores baixos de TCWV (mediana do erro menos valor real é menor que 1 K para valores menores que 30 cm de TCWV). Os autores também compararam  $T_S$  obtida dos Landsat com  $T_S$  *in situ* derivados de radiância de doze estações, onde se obtiveram baixos erros e vieses, com valores de raiz do erro quadrático médio em torno de 2°C.

É importante frisar que este algoritmo não é a única forma de se computar  $T_S$  por sensoriamento remoto, e todas elas carregam incerteza na estimativa. Essa incerteza surge não só pela incerteza de  $T_b$ , medido por satélite, mas também pelas incertezas das estimativas de emissividade da superfície e vapor na atmosfera. Variáveis que aqui não foram consideradas mas que também influenciam  $T_S$ , como o relevo, da mesma forma podem causar impacto na incerteza.

### 3.4 MODELOS

#### 3.4.1 Técnicas Estatísticas

O Coeficiente de Pearson ( $r$ ) é uma medição do grau de correlação linear entre duas variáveis, e assume valores adimensionais entre  $-1$  e  $1$  (DODGE, 2008). Dizer que duas variáveis são independentes entre si implica dizer que  $r$  é estatisticamente igual a zero.

O Método dos Mínimos Quadrados (MMQ) é uma técnica onde se estima parâmetros de uma função de tal forma que essa função represente da melhor forma possível a relação entre duas ou mais variáveis, fazendo com que a soma dos quadrados das diferenças entre o valor observado e o valor dado pela função seja mínimo (DODGE, 2008). É utilizado especialmente em Regressão Linear (ajuste linear entre duas variáveis).

#### 3.4.2 Fotoperíodo

O fotoperíodo ( $N$ ), ou duração do dia, é o tempo que o Sol fica acima do horizonte durante o dia e está intimamente relacionado com as estações do ano. Possui característica oscilatória. Fora da linha do Equador, um fotoperíodo abaixo de 12 horas se encontra no outono e inverno astronômicos (no geral mais frios), já um maior se encontra na primavera e verão astronômicos (no geral mais quentes). Para uma mesma latitude ( $\phi$ ),  $N$  varia apenas com a declinação solar ( $\delta$ ), que é o ângulo entre os raios solares e o plano do Equador. Segundo

Yamasoe e Corrêa (2016), o fotoperíodo pode ser calculado pelas equações (5) e (6).

$$N = 2H_0 \frac{12}{\pi} \quad (5)$$

$$H_0 = \arccos(-\operatorname{tg} \phi \operatorname{tg} \delta) \quad (6)$$

onde  $H_0$  é o ângulo horário do nascer do Sol e o fator  $\frac{12}{\pi}$  serve para converter radianos em horas. Segundo Paltridge e Platt (1976) é possível calcular  $\delta$  segundo função empírica dada pelas equações (7) e (8).

$$\begin{aligned} \delta = & 0,006\,918 - 0,399\,912 \cos(\theta_0) + 0,070\,257 \operatorname{sen}(\theta_0) + \\ & -0,006\,758 \cos(2\theta_0) + 0,000\,907 \operatorname{sen}(2\theta_0) + \\ & -0,002\,697 \cos(3\theta_0) + 0,001\,480 \operatorname{sen}(3\theta_0) \end{aligned} \quad (7)$$

$$\theta_0 = 2\pi \frac{(\text{NDA} - 1)}{365} \quad (8)$$

onde NDA é o número do dia do ano, ou seja, numera os dias de 1 a 365 começando em 1 de janeiro, e  $\theta_0$  é uma variável auxiliar na equação que converte NDA para radianos.

### 3.5 MÉTODOS

As posições das estações meteorológicas são fornecidas em latitude e longitude. Como a resolução espacial dos satélites Landsat é muito elevada, qualquer erro na posição colocará a estação em um pixel errado. Dessa forma, a posição das estações foi checada utilizando ferramentas de mapas disponíveis online (GEE e *Google Street View*), exceto quando não haviam imagens disponíveis para o local.

Estimativas de  $T_g$  foram calculadas no GEE. Para evitar contaminação por nuvens, pixels de nuvem e sombra de nuvem foram removidos utilizando máscaras do algoritmo *CFMask* (essas informações acompanham os dados Landsat). Como *CFMask* pode cometer erros, excluem-se também dias em que a RMSP esteja com uma cobertura maior que 50% de nuvens, o que implicaria numa maior chance de um pixel classificado como sem nuvem ser um falso negativo. Outra informação de qualidade contida nos dados Landsat é a saturação do pixel (i.e, a radiância que atingiu o sensor é maior ou igual ao máximo que o sensor consegue registrar) em cada banda. Todos os pixels saturados nas bandas utilizadas foram descartados.

Como descrito,  $N$  é uma variável sazonal que é maior que 12 horas para a primavera e verão e é menor que 12 para o outono e inverno. Toda análise entre variáveis foi feita separando

os dados nesses dois grupos.

### 3.5.1 Área de influência de $T_S$ sobre $T_a$

Uma das ferramentas disponíveis no GEE são os *kernels*. Com eles, é possível calcular qual é a média de  $T_S$  em torno de um dado ponto e dentro de um dado raio (para pixels incluídos parcialmente nesta circunferência, é feita a média ponderada em relação à área incluída). Calculou-se a média de  $T_S$  em 47 diferentes raios, indo de 15 a 70 000 m, com espaçamento entre pontos variando de forma a caracterizar melhor as relações estudadas. O principal objetivo aqui é encontrar em quais raios a correlação entre  $T_S$  e  $T_a$  é maximizada. Para que a amostra dentro do raio seja significativa, somente utilizaram-se médias onde havia menos de 50 % de cobertura de nuvem dentro da circunferência. E para consistência entre os raios, sempre que uma circunferência possuir mais de 50 % de cobertura de nuvens e for excluída, todas as demais circunferências de mesma estação meteorológica e dia também serão. Dessa forma, cada circunferência de uma mesma estação meteorológica possuirá o mesmo número de dados que as demais. Com as devidas correções, as posições das estações foram importadas para o GEE e em seguida se exportou todos os dados presentes nessas localizações.

Esses dados passaram por uma rotina em linguagem R que associa os dados de  $T_S$  com  $T_a$  de mesma localização, dia e horário. Realizou-se uma interpolação linear para que  $T_a$  se encontre no mesmo horário da passagem dos Landsat. Uma segunda rotina calculou o Coeficiente de Pearson ( $r$ ) de  $T_a$  com  $T_S$  para cada raio e estação meteorológica, de onde se elaboraram gráficos de  $r$  por raio. Esses gráficos foram agrupados visualmente de acordo com suas características. Gráficos de dispersão de  $T_a$  e  $T_S$  para algumas estações foram feitos de forma a ilustrar a linearidade, assim como ilustrar a melhora na correlação que calcular  $T_S$  em diferentes raios pode fornecer.

Por fim, uma estação meteorológica foi selecionada para se analisar a relação entre o raio de máxima correlação e as características do entorno da estação, utilizando uma composição RGB em cores naturais a partir das bandas do vermelho, verde e azul do satélite Landsat 8.

## 4 RESULTADOS E ANÁLISES

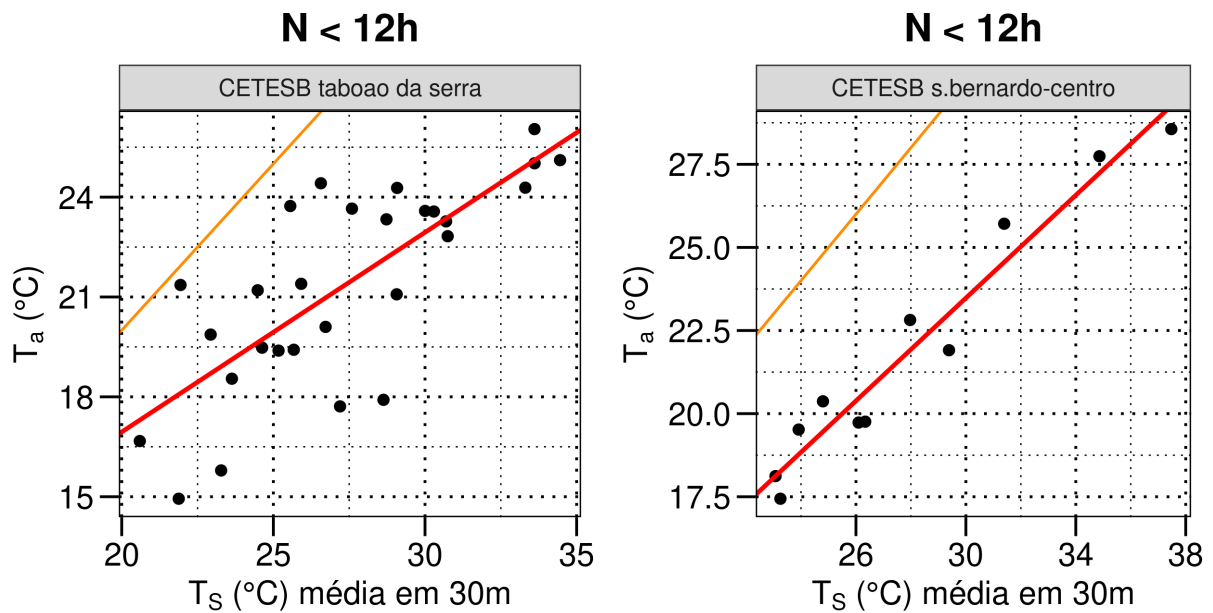
### 4.1 Distribuição entre $T_a$ e $T_S$

Como já descrito, neste trabalho se utiliza valores de  $T_a$  somente próximos das 10 horas no horário solar local e somente quando é possível obter  $T_S$  por sensoriamento remoto em uma grande área da RMSP (ou seja, quando há poucas nuvens e quando ocorre a revisita dos satélites,

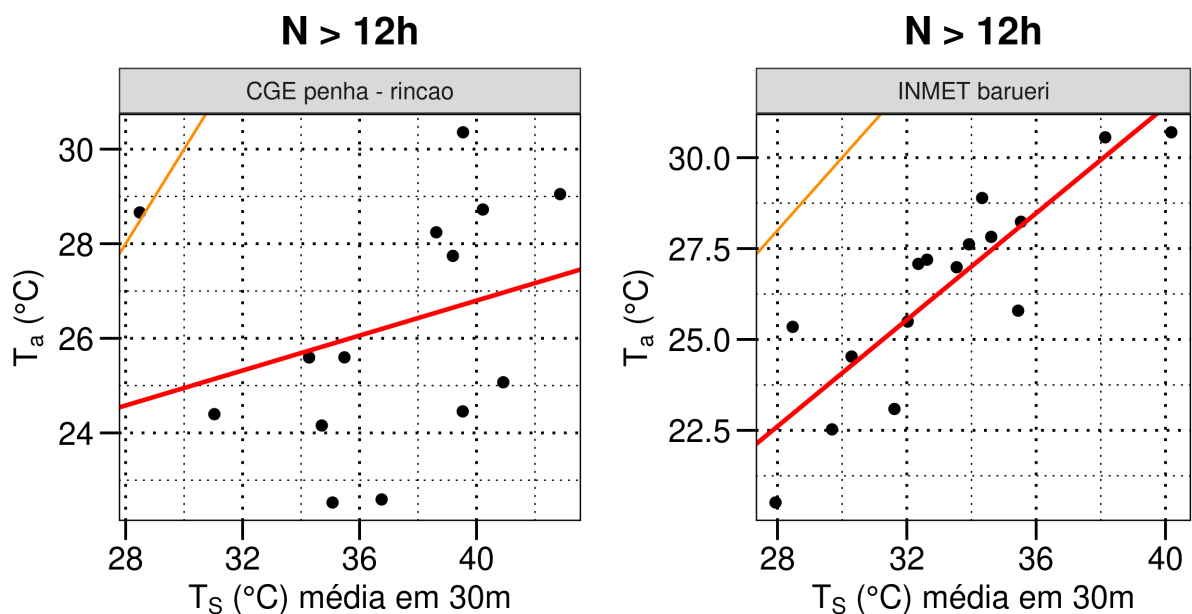
a cada 16 dias). Dessa forma, grande parte das interações entre  $T_a$  e  $T_S$  não serão caracterizados aqui (por exemplo, interações durante frentes frias, que deixam grande parte do céu nublado, ou também eventos atmosféricos que ocorrem com baixa frequência), embora os eventos aqui caracterizados ainda sejam relevantes. Como referência, nas Figuras B1 e B2 estão os histogramas de todos os dados de  $T_a$  (para todos os horários e dias com dados disponíveis) no período de estudo, e nas Figuras B3 e B4 estão gráficos *boxplot* dos mesmos dados. Com eles é possível visualizar quais estações são na média mais quente ou mais frias e também analisar qual o range de dados que foi utilizado neste estudo em relação aos dados originais. Como removeram-se os dias com nuvens, no geral mais frios que os dias sem nuvens (por nestes haver mais incidência solar), os dados de  $T_a$  efetivamente utilizados não incluirão os dados mais frios, como se observa nos gráficos mostrados ao longo do texto.

$T_a$  e  $T_S$  em um mesmo lugar aproximadamente às 10 h no horário solar local se relacionam de forma linear na maioria dos casos, como se observa nos exemplos das Figuras 2 e 3 (todas as estações utilizadas podem ser observadas nas Figuras C1 e C2). As Figuras 2 e 3 representam as estações meteorológicas de menor e maior correlação dentre as utilizadas, e para um fotoperíodo menor e maior que 12 h. Nessas figuras, utilizou-se o valor médio de  $T_S$  em um raio de 30 m ao redor de cada estação (a média é ponderada de acordo com a área do pixel do satélite incluso na circunferência), que será muito próximo ao valor do pixel sobre a estação meteorológica na resolução do satélite (120 m e 100 m reamostrados para 30 m, respectivamente para os satélites Landsat 5 e 8), e por isso é aqui utilizado como valor de referência em relação à resolução original do satélite. Aqui já se observa que a correlação é melhor para um fotoperíodo menor, e que ela varia para diferentes lugares (principalmente com um fotoperíodo maior, pois a diferença entre a correlação máxima e mínima é muito maior nesse caso). Além disso, dentre as estações com 10 ou mais dados utilizados, todas as estações obtiveram um p-valor menor que 0,05 para  $N < 12$  h (ou seja, o coeficiente angular é diferente de zero com uma confiança de pelo menos 95 % em todas as estações), enquanto que para  $N > 12$  h, 31 % (14 de 45 estações) obtiveram um p-valor maior. Isso implica dizer que a linearidade entre  $T_a$  e  $T_S$  médio em um raio de 30 m é bem assegurada para um fotoperíodo baixo, já para um alto, esse fato depende do lugar. Por fim, nas mesmas figuras ainda se observa que  $T_S$  é praticamente sempre maior que  $T_a$  ou pelo menos as duas são muito próximas. Isso ocorre pois nesse horário, a superfície se encontra sob aquecimento radiativo, e existe uma inércia na transferência de calor sensível da superfície para o ar na altura dos termômetros.

Na Figura 4 há um histograma com as correlações de todas as estações com 10 ou mais dados utilizados, e se observa novamente, agora em maiores detalhes, que a correlação para um fotoperíodo menor que 12 h é maior que para um maior que 12, e a variância da distribuição é

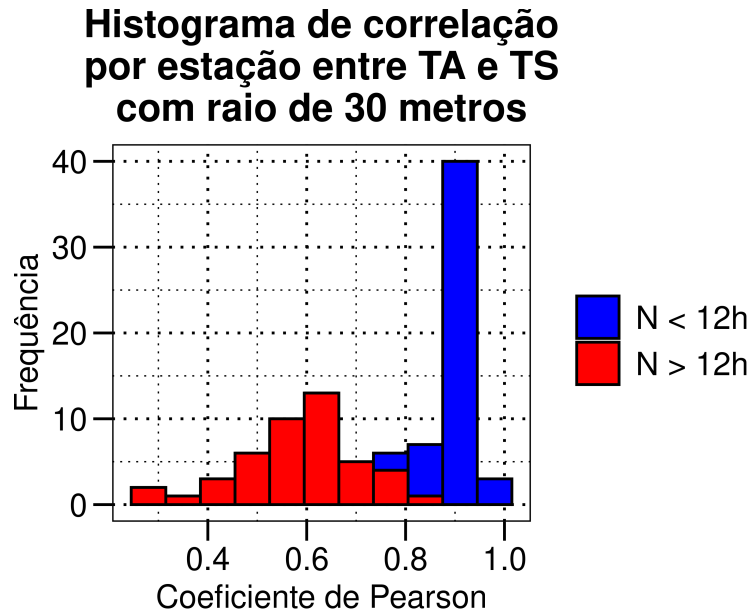


**Figura 2:** Dispersão entre  $T_a$  e  $T_S$  média em raio de 30 m, aproximadamente às 10 h no horário solar local, com fotoperíodo **menor** que 12 h e para as estações meteorológicas com menor e maior correlação, respectivamente a primeira e segunda imagem (considerando apenas estações com 10 ou mais dados de  $T_a$  e  $T_S$  disponíveis). A primeira possui  $r = 0,757$  e a segunda,  $0,974$ . Os p-valores dos coeficientes angulares são respectivamente  $3,1 \cdot 10^{-6}$  e  $4,2 \cdot 10^{-7}$ . Linha laranja é a reta 1:1 e vermelha é o ajuste linear.



**Figura 3:** O mesmo que a Figura 2, mas com fotoperíodo **maior** que 12 h. A primeira imagem possui  $r = 0,291$  e a segunda,  $0,867$ , e os p-valores dos coeficientes angulares são respectivamente  $0,31$  e  $1,4 \cdot 10^{-5}$ .





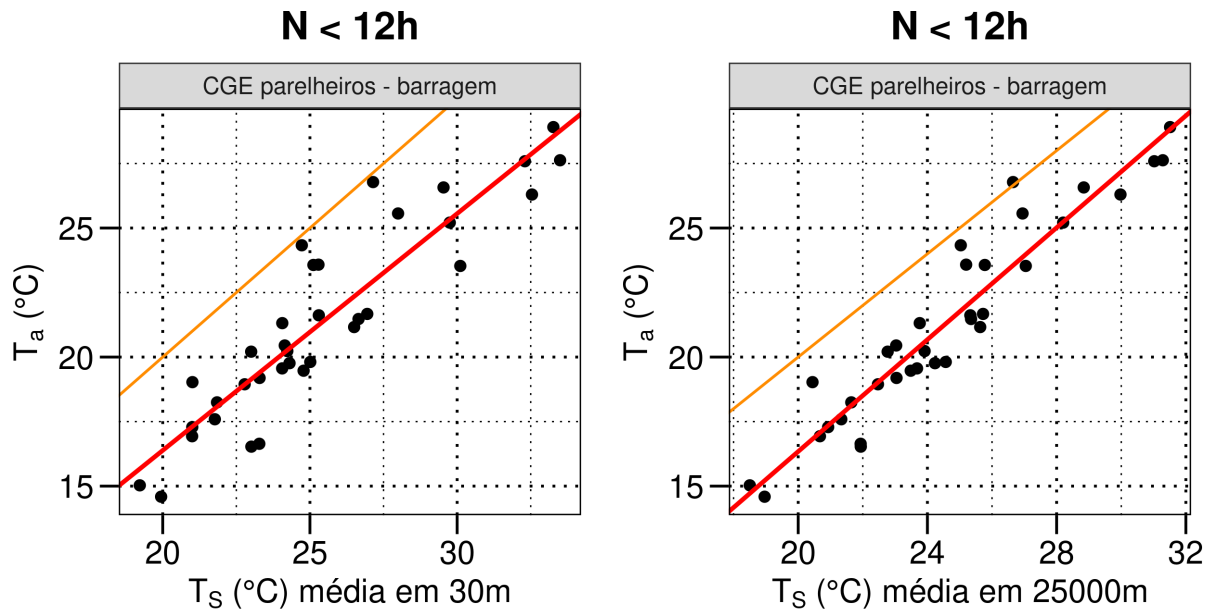
**Figura 4:** Histograma de correlações entre  $T_a$  e  $T_S$  médio em raio de 30 m para cada estação meteorológica utilizando apenas estações com 10 ou mais dados. Separaram-se os dados por fotoperíodo.

maior para um fotoperíodo maior.

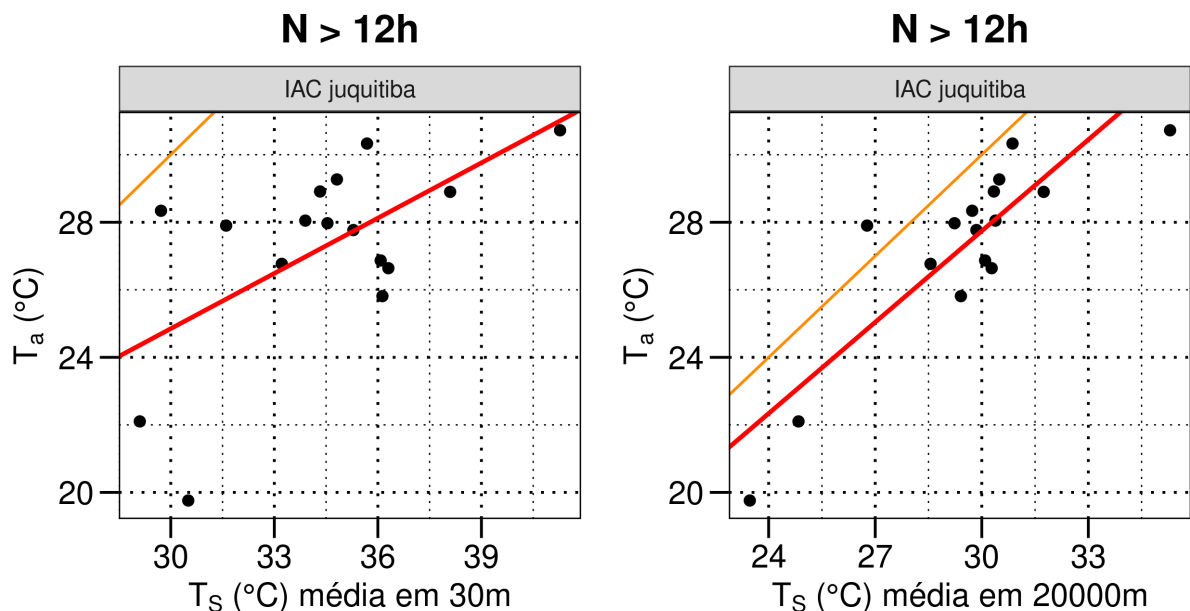
#### 4.2 Influência do Entorno na Correlação

Ao se utilizar um valor de  $T_S$  médio em circunferências de diferentes raios (além do de 30 m, utilizado anteriormente), a correlação entre  $T_a$  e  $T_S$  pode ser melhorada. Para muitas estações, realizar esse procedimento traz poucas mudanças, mas em alguns casos a correlação melhora de forma significativa, como mostram os exemplos das Figuras 5 e 6, que trazem gráficos de distribuição para duas estações, uma com  $N < 12$  h e outra para  $N > 12$  h, entre  $T_a$  e  $T_S$  média em um raio de 30 m e  $T_S$  média em um raio cujo a correlação foi máxima. Percebe-se que os pontos ficam muito mais próximos do ajuste no segundo caso. Além disso, para  $N > 12$  h e para estações com 10 ou mais dados, a porcentagem de estações que possuem p-valor menor que 0,05 aumentou de 31 % para 84 % (38 de 45), um aumento muito significativo (para  $N < 12$  h, essa fração permaneceu em 100%), ou seja, muitas estações que não possuem boa linearidade entre  $T_a$  e  $T_S$  médio em 30 m passam a ter boa linearidade ao considerar raios maiores. Nos exemplos das figuras, o raio que maximiza a correlação é da ordem de 20 km, que pertence à mesoescala na meteorologia. É possível que isso ocorra em grande parte pela advecção do ar, fazendo-o ter contato com a superfície em diversos locais antes de chegar ao ponto onde se encontra a estação meteorológica, e mistura e trocas de calor entre parcelas de ar próximas.

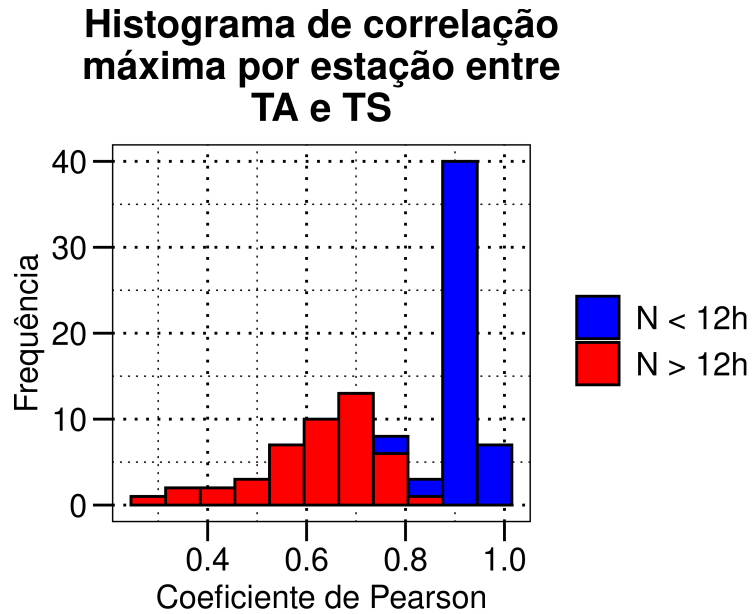
Substituindo a correlação utilizando  $T_S$  média em 30 m pela correlação máxima para cada estação, obtém-se o histograma da Figura 7. Comparando com o histograma da Figura 4,



**Figura 5:** Dispersão entre  $T_a$  e  $T_s$  média em raio de 30 m (primeira imagem) e em um raio onde a correlação se maximiza (segunda imagem), para a estação de Parelheiros (gestor CGE), com fotoperíodo **menor** que 12 h. A primeira possui  $r = 0,910$  e a segunda,  $0,950$ . Os p-valores dos coeficientes angulares são respectivamente  $3,5 \cdot 10^{-14}$  e  $3,3 \cdot 10^{-18}$ . Linha laranja é a reta 1:1 e vermelha é o ajuste linear.



**Figura 6:** O mesmo que a Figura 5, mas para a estação de Juitiba (gestor IAC) e com fotoperíodo **maior** que 12 h. A primeira imagem possui  $r = 0,607$  e a segunda,  $0,867$ , e os p-valores dos coeficientes angulares são respectivamente  $0,0127$  e  $1,4 \cdot 10^{-5}$ .



**Figura 7:** Histograma de correlações entre  $T_a$  e  $T_S$  médio no raio que maximiza a correlação, para cada estação meteorológica e utilizando apenas estações com mais de 10 dados. Separaram-se os dados por fotoperíodo.

para  $N > 12$  h nota-se que os extremos são pouco afetados pela mudança, porém a classe de maior frequência se desloca para valores maiores. Já para  $N < 12$  h, o pico permanece inalterado, assim como o extremo inferior, e a segunda classe diminui enquanto a classe superior aumenta. Para ambos os grupos de fotoperíodo, conclui-se que há estações meteorológicas em que é possível aumentar a correlação ao se considerar o entorno, enquanto em outras, não é.

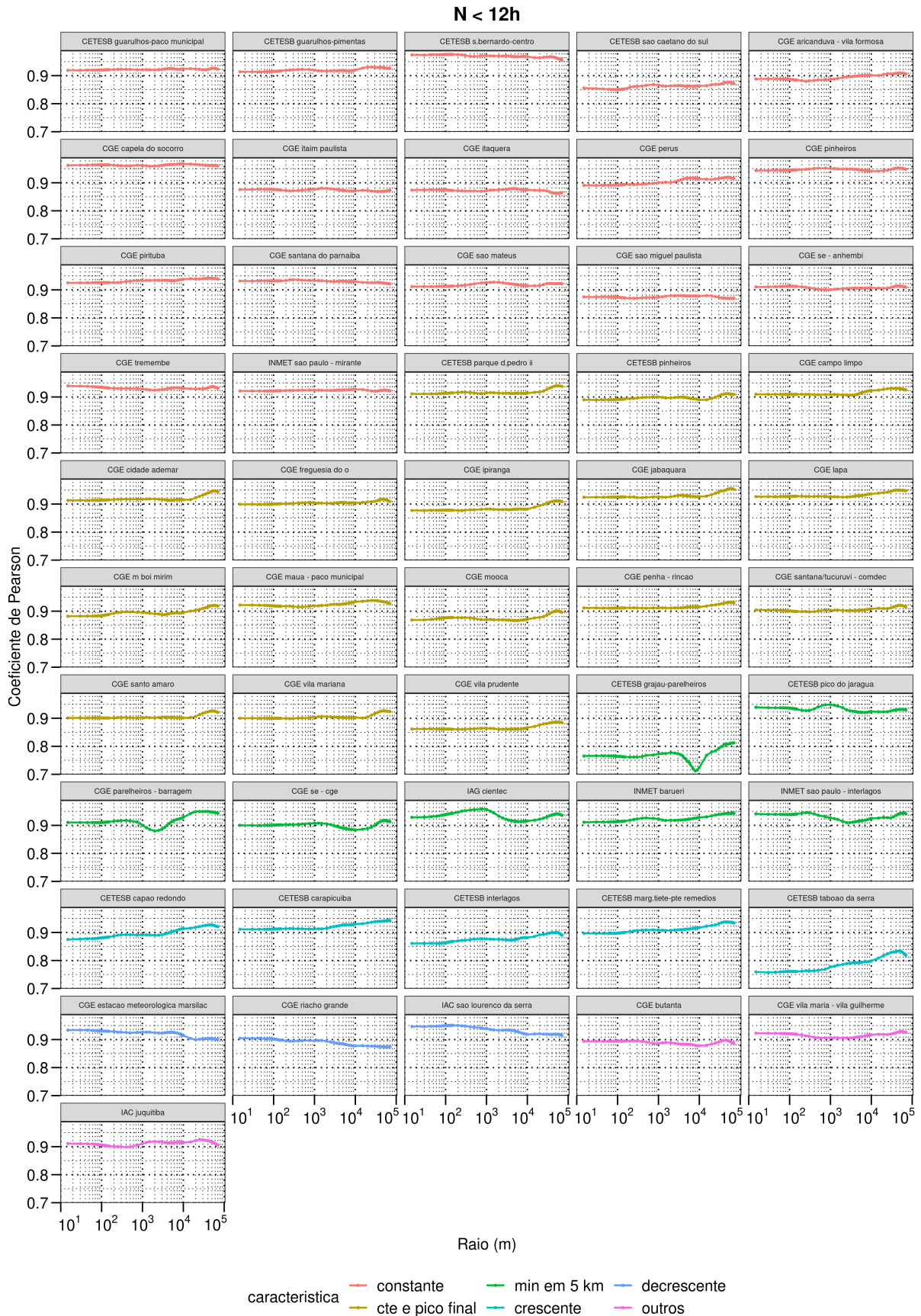
Nas Figuras 8 e 9 estão representadas as correlações para cada estação meteorológica em função do raio onde se calculou a média de  $T_S$ , respectivamente para  $N < 12$  h e para  $N > 12$  h. As estações estão divididas em diferentes grupos, de acordo com o comportamento da curva característica (embora possa ocorrer casos onde a estação possui características de mais de um grupo). Analisando primeiro o caso onde  $N < 12$  h, nota-se que uma parcela expressiva das estações possui uma correlação praticamente constante ao se variar o raio. Em outras palavras, considerar o entorno possui pouco impacto na correlação de  $T_a$  e  $T_S$ . Isso não significa necessariamente que não existe um raio de maior influência de  $T_S$  sobre  $T_a$ , já que a composição majoritária da superfície (e.g., residências) pode ser a mesma em uma grande área ao redor da estação, fazendo  $T_S$  ser pouco alterada (aqui essa hipótese não foi checada). O segundo maior grupo é similar ao primeiro, exceto no final do eixo x, em raios da ordem de  $10^4$  m ou maiores, onde ocorre um aumento na correlação, e no geral aparece um ponto de máximo próximo de  $5 \cdot 10^4$  m. Essa é a distância aproximada entre o centro da cidade de São Paulo (para referência utilizando a Figura 1, fica em torno de  $23.55^\circ\text{S}$ ,  $46.63^\circ\text{W}$ ) e o mar. Dessa forma, é possível que a diminuição na correlação após o pico ocorra pela presença do mar, porém não se

sabe o que pode causar o aumento em primeiro lugar. Como na estação da Figura 5, a ordem de grandeza do máximo pertence à mesoescala, e deve estar relacionado à advecção e misturas entre parcelas de ar.

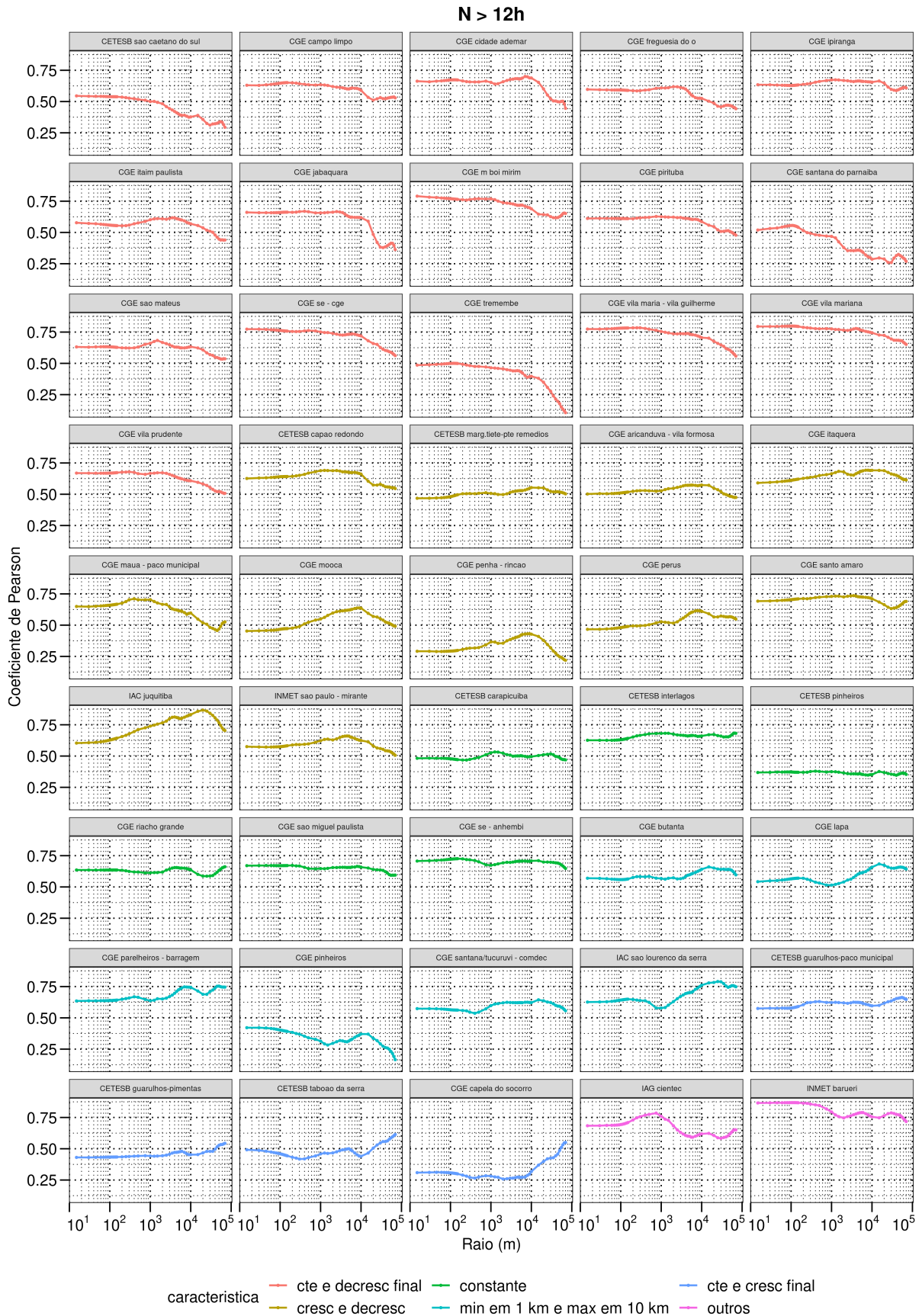
O terceiro grupo possui aparência mais heterogênea e é caracterizado por um ponto de mínimo perto de 5 km (entre 2 e 10 km). Em outras palavras, a correlação atinge um máximo local em pelo menos dois raios no domínio, um da ordem de 1 km ou menor e outro da ordem de 10 km ou maior. A discussão para picos em raios maiores é similar à feita para o grupo anterior, já para picos em raios menores, estes pertencem à microescala na meteorologia, a qual trocas de energia entre a superfície e o ar fazem parte. Dessa forma, essa circunferência de correlação máxima deve estar relacionada com a área da superfície que influencia diretamente a temperatura do ar, sem grandes contribuições de mistura e trocas de energia entre parcelas de ar.

O quarto e o quinto grupo representam respectivamente uma correlação que é crescente e decrescente com o raio. Algumas estações em que é crescente também formam um pico no final do eixo  $x$ , ao redor de 50 km, como em grupos anteriores. O grupo em que é crescente possui análise análoga ao comportamento constante com um pico no final. Já o que é decrescente, possui máximo da ordem de 100 m, que também se encontra na microescala. Os processos aqui possuem ainda menos contribuição de trocas entre parcelas de ar, e as trocas de calor entre ar e superfície num mesmo local são ainda mais significativas. Por fim, há também o grupo caracterizado como “outros”, que possuem atributos presentes em mais de um grupo, a ponto de ser difícil definir um grupo ao qual esta se encaixa melhor.

Analisando agora as estações com  $N > 12$  h, percebe-se que as curvas características no geral são diferentes de para  $N < 12$  h. O grupo mais comum possui uma correlação praticamente constante com o raio, mas decrescente para raios elevados. O segundo comportamento mais comum é similar ao primeiro, é decrescente em raios elevados, mas ao invés de constante, é crescente em raios menores. Nesse grupo, na maioria dos casos o pico ocorre ao redor de 10 km, mas também ocorrem casos em 1 km. Em outras palavras, na maioria dos casos a correlação é máxima dentro da mesoescala, mas ocorrem casos na microescala. No próximo grupo há estações de comportamento mais heterogêneo caracterizadas por um mínimo em 1 km e um máximo em 10 km. Dessa forma, há também um máximo da ordem de 100 m e logo esse grupo também possui máximos tanto na meso quanto na microescala. Mesmo assim, no geral o pico na macroescala é maior do que na micro. Em seguida, há o grupo onde a correlação é praticamente constante, que foi o grupo dominante em  $N < 12$  h. O grupo seguinte é crescente em raios elevados, e no geral é constante nos demais. Para algumas estações, esse grupo é similar ao do  $N < 12$  h onde a correlação é constante, com um pico no final. E por fim, há o grupo “outros”,



**Figura 8:** Correlação entre  $T_a$  e  $T_S$  médio em diferentes raios ao redor das estações meteorológicas, para um fotoperíodo **menor** que 12 h. Estações foram agrupadas em diferentes categorias, de acordo com sua curva característica.



**Figura 9:** Correlação entre  $T_a$  e  $T_S$  médio em diferentes raios ao redor das estações meteorológicas, para um fotoperíodo **maior** que 12 h. Estações foram agrupadas em diferentes categorias, de acordo com sua curva característica.

que não se encaixaram bem nos demais grupos. Nas duas estações desse grupo o máximo ocorre na microescala, uma ao redor de 100 m e outra ao redor de 1 km.

### 4.3 Comparação Preliminar Entre Composição do Entorno e Correlação



**Figura 10:** Composição RGB em cores naturais do Landsat 8 no dia 28/03/2020. Ponto azul demarca a localização da estação meteorológica do Cientec (IAG), 23,651 242°S 46,622 424°W, localizada no Parque do Estado, e circunferência demarca um raio de 750 m ao redor da estação.

Nas Figuras 8 e 9 foi observado que há estações com um pico na correlação em raios da ordem de 1 km. Um desses exemplos é a estação do IAG Cientec, que possui um máximo na correlação em 750 m, tanto para  $N$  maior quanto menor que 12 h. Essa estação está localizada em uma reserva ambiental dentro da área urbana da RMSP chamada Parque do Estado. Na Figura 10, há uma composição RGB em cores naturais do satélite Landsat 8 da área do parque e seu entorno, assim como a localização da estação e o raio de 750 m ao redor da mesma. É notável que a máxima correlação ocorre quando a média de  $T_S$  é calculada em uma circunferência que praticamente tangencia a área urbana fora do parque. Ao aumentar a circunferência, a  $T_S$  da área urbana começa a ser incluída na média e a correlação cai. É possível que as estações que possuem maiores variações na correlação com o raio também possuam esse tipo de fronteira em suas proximidades, ao mesmo tempo em que, como já discutido, as estações onde há pouca variação na correlação devem possuir superfícies na média mais uniformes, de forma que as áreas de maior influência não possam ser determinadas, embora uma análise mais detalhada dessas superfícies seja necessária.

## 5 CONCLUSÕES E RECOMENDAÇÕES

A atmosfera se relaciona de formas complexas com a superfície, e este estudo ilustra como a correlação entre  $T_a$  e  $T_S$  muda para diferentes locais relativamente próximos e para diferentes períodos do ano. Dentre as estações meteorológicas utilizadas, a correlação em questão é maior para  $N < 12$  h do que para maior e varia muito mais espacialmente para  $N > 12$  h do que para menor.

Considerar valores de  $T_S$  do entorno da estação meteorológica consegue melhorar a correlação com  $T_a$ , embora não para todas as estações. Mesmo assim, a melhora é significativa. No geral, o raio das circunferências que mais beneficiam a correlação entre  $T_S$  e  $T_a$  varia bastante de estação para estação, mas comportamentos comuns foram encontrados. A escala dos raios de máxima correlação também atingem tanto a microescala quanto a mesoescala da meteorologia, e dessa forma, é possível que fenômenos atmosféricos de ambas as escalas estejam influenciando no valor final de  $T_a$ . Na microescala, por exemplo, ocorrem as trocas de calor entre a superfície e o ar, já na mesoescala ocorrem transportes de parcelas de ar e mistura e trocas de calor entre elas.

Por fim, analisando os entornos de uma estação meteorológica específica, conjectura-se que muitas das estações onde a correlação não varia com o raio se comportam dessa forma pois as características do entorno da estação são na média uniformes, mas ainda é necessário analisar a composição do entorno de mais estações para chegar em conclusões melhores. Uma próxima etapa para este estudo seria verificar se há relações entre as estações meteorológicas de curvas características similares e o entorno delas, e dessa forma conseguir estabelecer em qual raio a máxima correlação ocorrerá baseado apenas nas informações de superfície. Com os resultados deste estudo ainda não é possível estabelecer um raio comum como recomendável para maximizar a correlação.

Dessa forma, para futuros estudos, recomenda-se entender que características da superfície impactam em qual raio ocorrerá a máxima correlação. Além disso, utilizar circunferências ainda é um método simplificado. Pode ainda ser melhorado se considerar, por exemplo, a velocidade e direção do vento, considerando que uma parcela de ar é influenciada pela superfície pela qual ela passou sobre, com um maior peso para as mais recentes. O método também não considera variações em variáveis como a altitude, que podem impactar  $T_a$  e  $T_S$  de formas diferentes. A influência de corpos d'água também não foi analisada. E não menos importante, aqui se estudou apenas dados das 10:00 no horário solar local. Sem dúvidas diferentes comportamentos devem ocorrer em outros horários. Com o avanço nos produtos por satélite, ao final do período deste trabalho se tornou disponível no GEE um novo produto de  $T_S$ , feito pela própria



USGS. Como este produto e o aqui utilizado são estimativas, eles devem ser diferentes em certos aspectos. Deve ser interessante comparar como os dois produtos se diferem e se um deles se relaciona melhor com  $T_a$  do que o outro. Como ainda não é possível estimar  $T_a$  diretamente por dados de sensoriamento remoto, ainda há muito a se compreender para que isso seja possível.

## REFERÊNCIAS

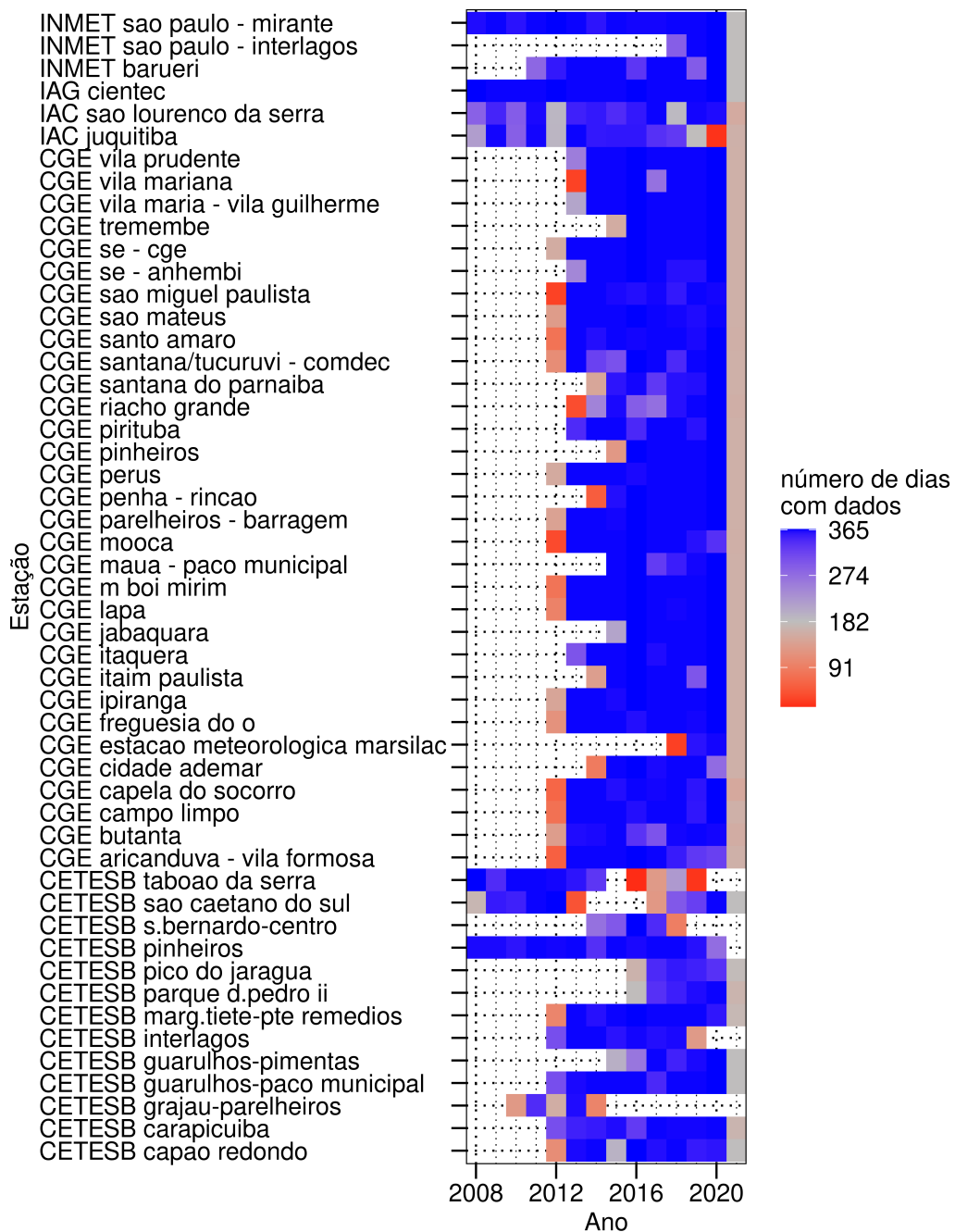
- CHENG, K.-S. et al. Assessing the effect of landcover changes on air temperature using remote sensing images—a pilot study in northern taiwan. **Landscape and Urban Planning**, Elsevier, v. 85, n. 2, p. 85–96, 2008.
- DEMOGRAPHIA. **Demographia World Urban Areas**: Built-up urban areas or urban agglomerations. 17. ed. 2021. Anual.
- DODGE, Y. **The Concise Encyclopedia of Statistics**. New York, NY: Springer, 2008. 616 p. ISBN 978-0-387-32833-1. Disponível em: <[https://doi.org/10.1007/978-0-387-32833-1\\_62](https://doi.org/10.1007/978-0-387-32833-1_62)>.
- ERMIDA, S. L. et al. Google earth engine open-source code for land surface temperature estimation from the landsat series. **Remote Sensing**, Multidisciplinary Digital Publishing Institute, v. 12, n. 9, p. 1471, 2020.
- FERREIRA, L. S. **Vegetação, temperatura de superfície e morfologia urbana: um retrato da região metropolitana de São Paulo**. 196 p. Doutorado em Arquitetura e Urbanismo — Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo, 2019.
- GOOD, E. J. et al. A spatiotemporal analysis of the relationship between near-surface air temperature and satellite land surface temperatures using 17 years of data from the ATSR series. **Journal of Geophysical Research: Atmospheres**, Wiley Online Library, v. 122, n. 17, p. 9185–9210, 2017.
- GORELICK, N. et al. Google earth engine: Planetary-scale geospatial analysis for everyone. **Remote sensing of Environment**, Elsevier, v. 202, p. 18–27, 2017.
- LANDSAT Satellite Missions. **USGS**. Disponível em: <<https://www.usgs.gov/core-science-systems/nli/landsat/landsat-satellite-missions>>. Acesso em: 13 nov. 2021.
- LI, H. et al. A new method to quantify surface urban heat island intensity. **Science of the total environment**, Elsevier, v. 624, p. 262–272, 2018.
- PALTRIDGE, G. W.; PLATT, C. M. R. In: \_\_\_\_\_. **Radiative Processes in Meteorology and Climatology**. [S.l.]: Elsevier Science, 1976. p. 60–63.
- PETTY, G. W. Components of the earth's energy budget. In: \_\_\_\_\_. **A first course in atmospheric radiation**. 2. ed. [S.l.]: Sundog Pub, 2006. cap. 1, p. 6–7.
- REBOITA, M. S. et al. Regimes de precipitação na américa do sul: uma revisão bibliográfica. **Revista brasileira de meteorologia**, SciELO Brasil, v. 25, n. 2, p. 185–204, 2010.
- SHENG, Y. et al. Quantifying the spatial and temporal relationship between air and land surface temperatures of different land-cover types in southeastern china. **International journal of remote sensing**, Taylor & Francis, v. 38, n. 4, p. 1114–1136, 2017.
- STULL, R. B. Satellites & radar. In: \_\_\_\_\_. **Practical meteorology: an algebra-based survey of atmospheric science**. [S.l.]: The University of British Columbia, 2015. cap. 8, p. 219–266.
- SZABÓ, S.; GÁCSI, Z.; BALÁZS, B. Specific features of NDVI, NDWI and MNDWI as reflected in land cover categories. **Landscape & Environment**, v. 10, n. 3-4, p. 194–202, 2016. ISSN 1789-4921.

YAMASOE, M. A.; CORRÊA, M. de P. In: \_\_\_\_\_. **Processos radiativos na atmosfera**. [S.l.]: Oficina de Textos, 2016. p. 44–53.

YAN, H. et al. Quantifying the impact of land cover composition on intra-urban air temperature variations at a mid-latitude city. **Plos one**, Public Library of Science, v. 9, n. 7, p. e102124, 2014.

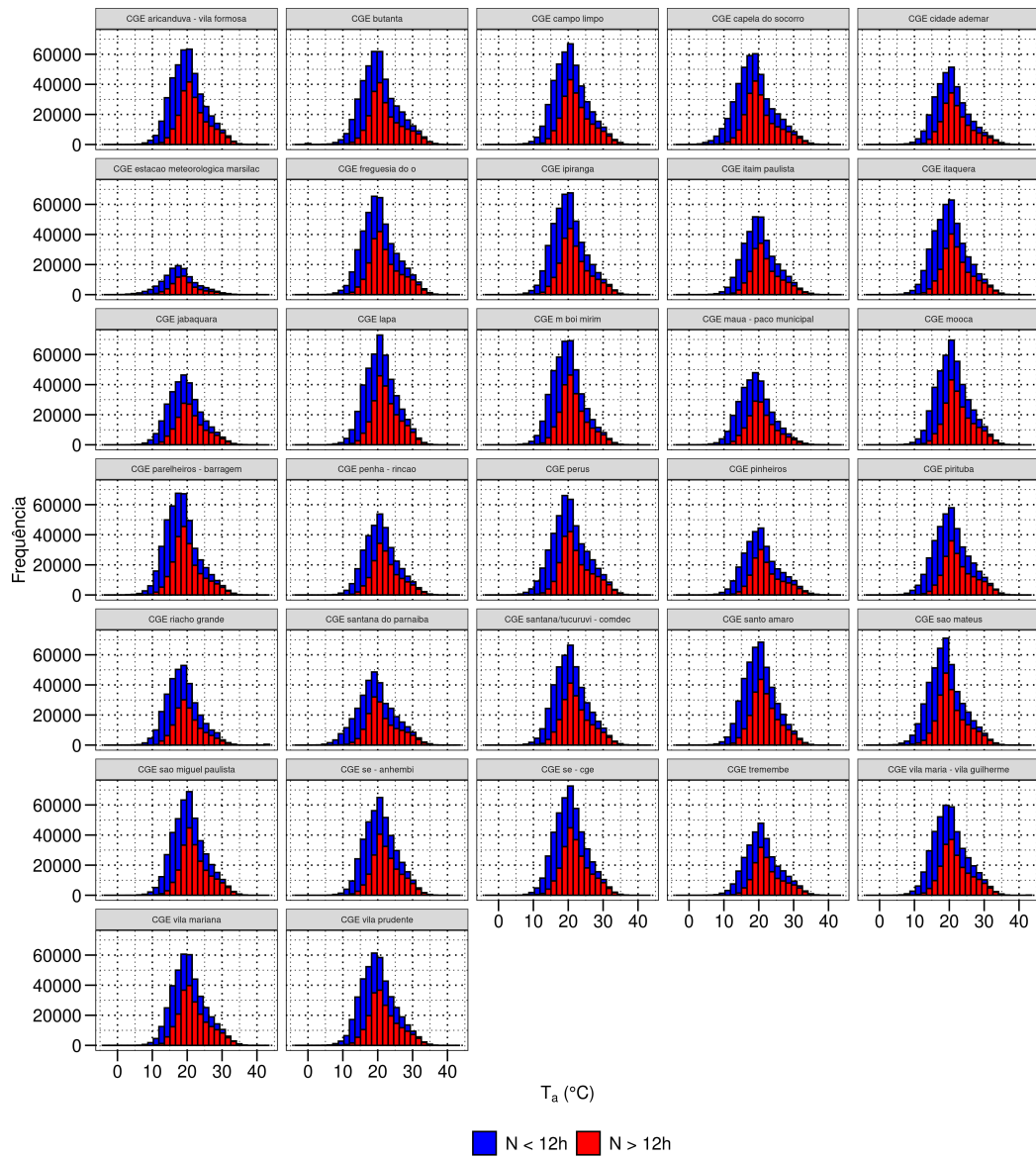
YOKOBORI, T.; OHTA, S. Effect of land cover on air temperatures involved in the development of an intra-urban heat island. **Climate Research**, v. 39, n. 1, p. 61–73, 2009.

## APÊNDICE A – Disponibilidade de dados de $T_a$ por estação

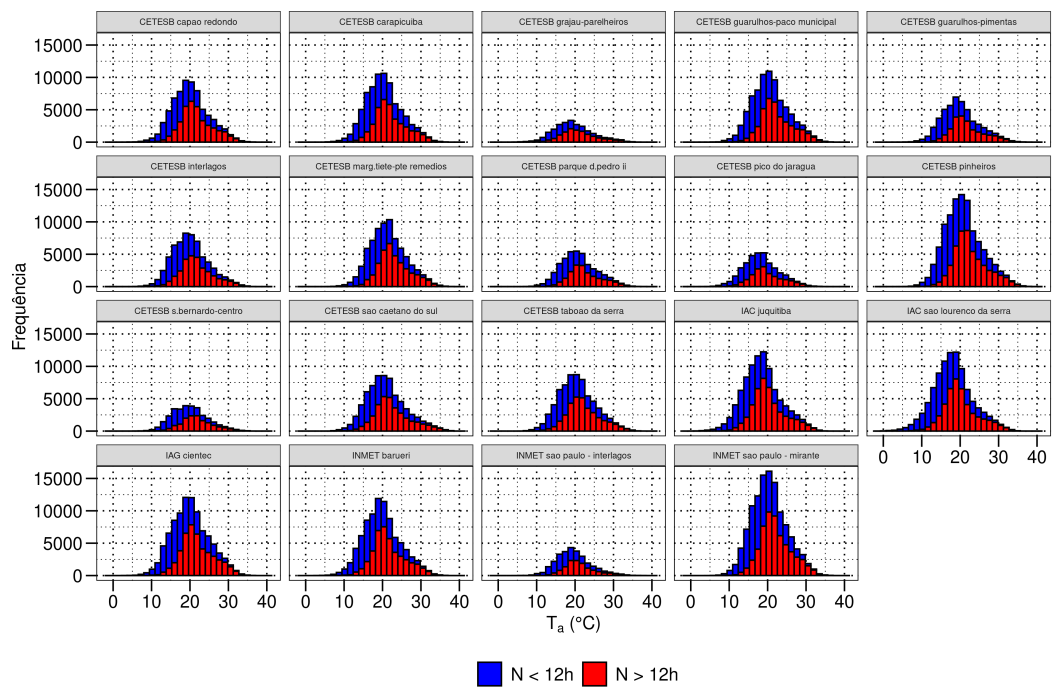


**Figura A1:** Número de dias no ano com dados de  $T_a$  disponíveis, por estação meteorológica.

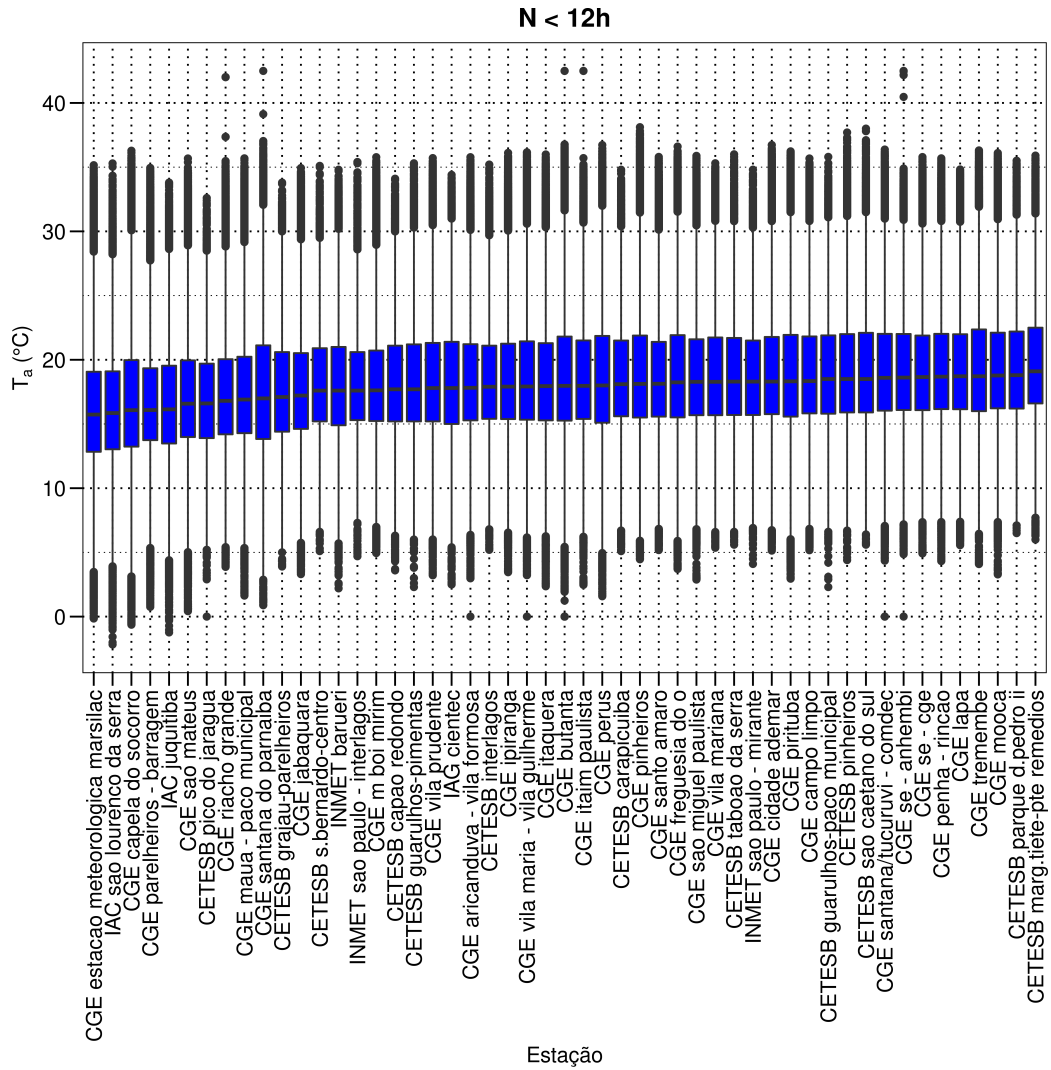
## APÊNDICE B – Distribuição de $T_a$ por estação



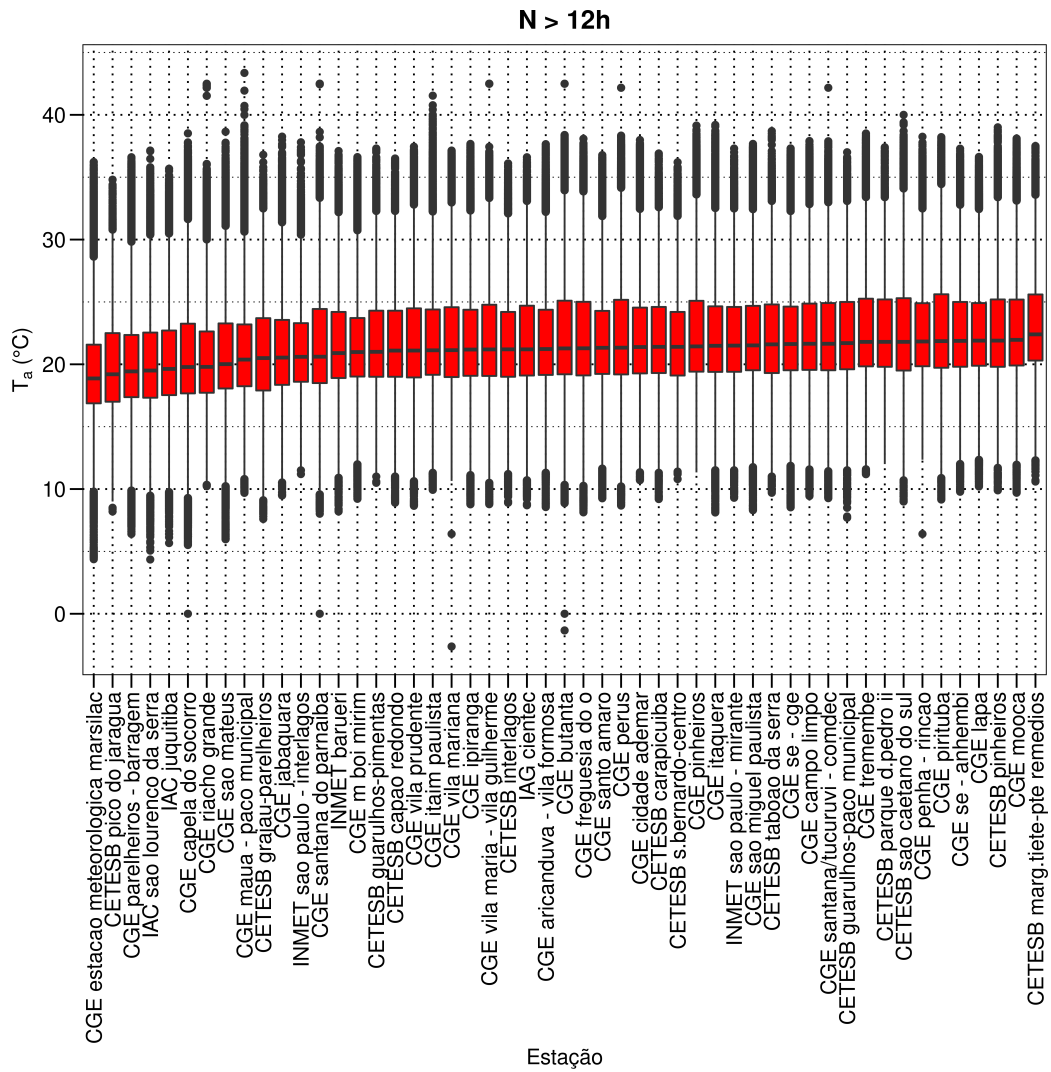
**Figura B1:** Histogramas de  $T_a$  por estação meteorológica para todos os horários, entre 2008 e junho de 2021 e para as estações do CGE. Os dados utilizados do CGE possuem resolução temporal de 20 minutos.



**Figura B2:** Histogramas de  $T_a$  por estação meteorológica para todos os horários, entre 2008 e junho de 2021 e para as estações da CETESB, IAC, IAG e INMET. Os dados dessas estações possuem resolução temporal de 1 hora.



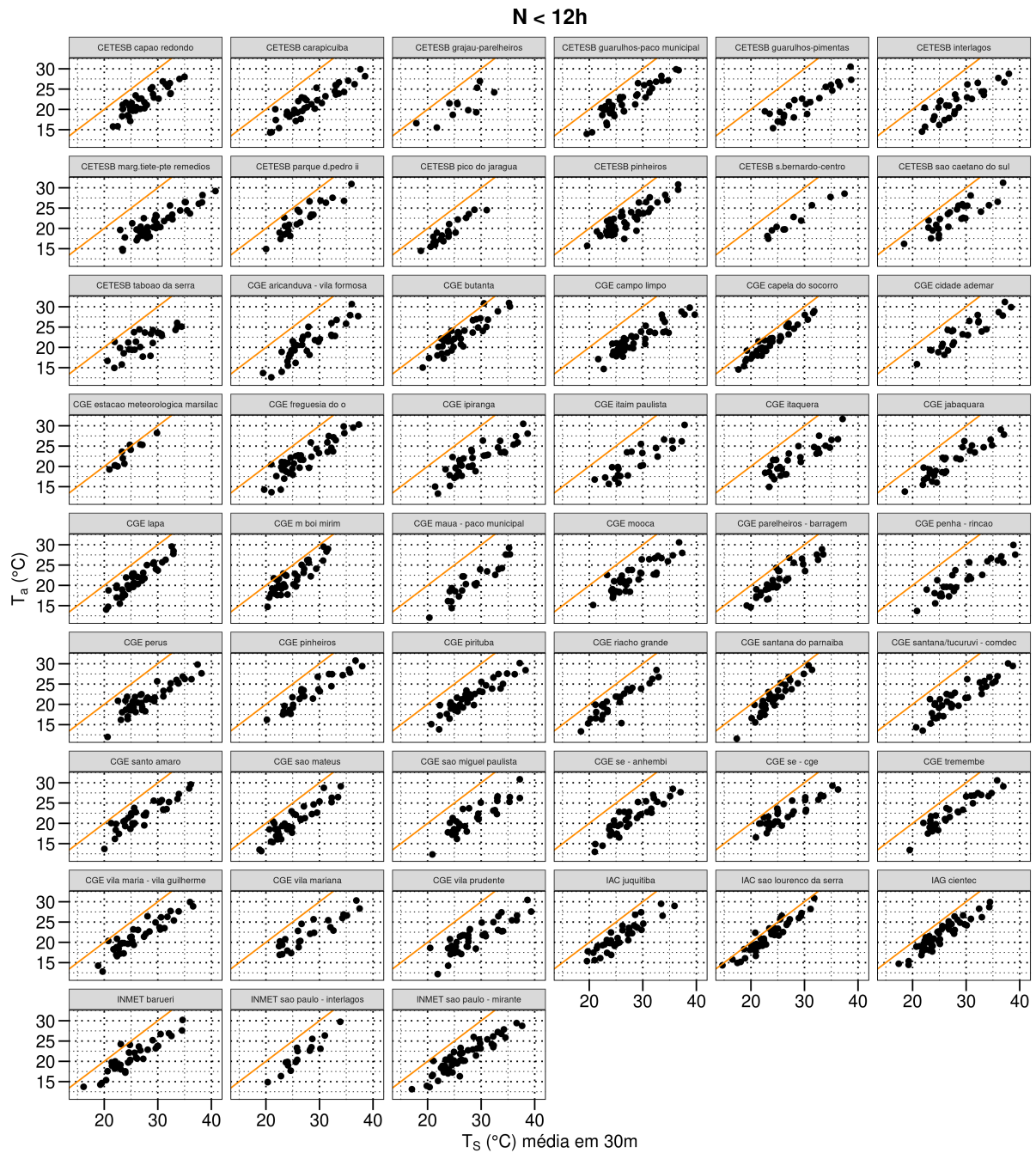
**Figura B3:** Gráficos *boxplot* de  $T_a$  por estação meteorológica para todos os horários, entre 2008 e junho de 2021 e para um fotoperíodo **menor** que 12 h.



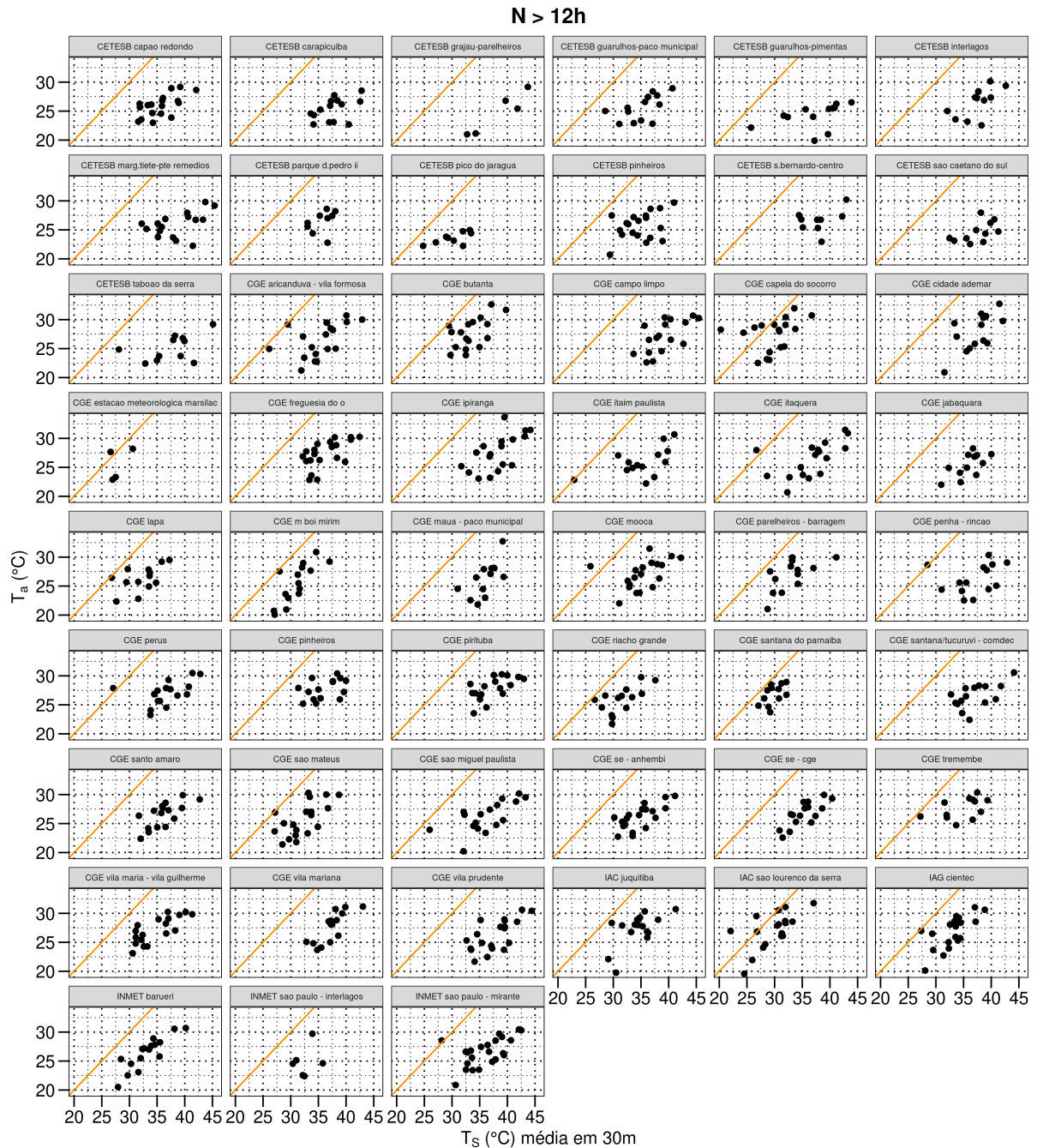
**Figura B4:** Gráficos *boxplot* de  $T_a$  por estação meteorológica para todos os horários, entre 2008 e junho de 2021 e para um fotoperíodo **maior** que 12 h.



## APÊNDICE C – Dispersão entre $T_a$ e $T_S$ para cada estação meteorológica



**Figura C1:** Dispersão entre  $T_a$  e  $T_S$  média em raio de 30 m, com fotoperíodo **menor** que 12 h e para cada estação meteorológica. Linha laranja é a reta 1:1.



**Figura C2:** Dispersão entre  $T_a$  e  $T_s$  média em raio de 30 m, com fotoperíodo **maior** que 12 h e para cada estação meteorológica. Linha laranja é a reta 1:1.