

Ministério do Planejamento, Orçamento e Gestão  
**Instituto Brasileiro de Geografia e Estatística – IBGE**  
Diretoria de Pesquisas  
Coordenação de Métodos e Qualidade

**Textos para discussão**  
**Diretoria de Pesquisas**  
**número 23**

# **Amostra Mestra para o Sistema Integrado de Pesquisas Domiciliares**

Marcos Paulo Soares de Freitas  
Maurício Franca Lila  
Rosemary Vallejo de Azevedo  
Giuseppe de Abreu Antonaci

Rio de Janeiro  
2007

**Instituto Brasileiro de Geografia e Estatística - IBGE**  
Av. Franklin Roosevelt, 166 - Centro - 20021-120 - Rio de Janeiro, RJ - Brasil

ISSN 1518-675X **Textos para discussão. Diretoria de Pesquisas**

Divulga estudos e outros trabalhos técnicos desenvolvidos pelo IBGE ou em conjunto com outras instituições, bem como resultantes de consultorias técnicas e traduções consideradas relevantes para disseminação pelo Instituto. A série está subdividida por unidade organizacional e os textos são de responsabilidade de cada área específica.

ISBN 978-85-240-3988-1

© IBGE. 2007

## **Impressão**

Gráfica Digital/Centro de Documentação e Disseminação de Informações - CDDI/IBGE, em 2007.

## **Capa**

Gerência de Criação/CDDI

---

Amostra mestra para o sistema integrado de pesquisas domiciliares /  
Marcos Paulo Soares de Freitas ... [et al.]. - Rio de Janeiro : IBGE,  
Coordenação de Métodos e Qualidade, 2007.  
p. – (Textos para discussão. Diretoria de Pesquisas, ISSN 1518-  
675X ; n. 23)

Inclui bibliografia.  
ISBN 978-85-240-3988-1

1. Levantamentos domiciliares – Brasil – Avaliação. 2. Amostragem  
(Estatística). 3. Domicílios – Pesquisas – Brasil. I. Freitas, Marcos Paulo  
Soares de. II. IBGE. Coordenação de Métodos e Qualidade. III. Série.

**Gerência de Biblioteca e Acervos Especiais**  
RJ/2007-31

CDU 314.6(81)  
DEM

---

Impresso no Brasil / Printed in Brazil

# Sumário

<b>APRESENTAÇÃO .....</b>	<b>5</b>
<b>INTRODUÇÃO .....</b>	<b>7</b>
<b>POPULAÇÃO ALVO E ABRANGÊNCIA GEOGRÁFICA.....</b>	<b>8</b>
<b>PLANO AMOSTRAL .....</b>	<b>8</b>
CADASTRO DE SELEÇÃO.....	8
DEFINIÇÃO DAS UNIDADES PRIMÁRIAS DE AMOSTRAGEM.....	9
ESTRATIFICAÇÃO .....	9
<i>Tamanho mínimo dos estratos .....</i>	<i>12</i>
<i>Estratificação por divisão administrativa.....</i>	<i>13</i>
<i>Estratificação geográfica e espacial .....</i>	<i>14</i>
<i>Estratificação por situação do setor.....</i>	<i>16</i>
<i>Estratificação estatística .....</i>	<i>18</i>
<b>TAMANHO DA AMOSTRA.....</b>	<b>19</b>
ESQUEMA DE ROTAÇÃO DA AMOSTRA DE DOMICÍLIOS NA PNADC .....	23
ALOCÇÃO DO TAMANHO DA AMOSTRA .....	23
<b>AVALIAÇÃO DA PRECISÃO DA ESTIMATIVA DE OUTROS PARÂMETROS E PARA OUTROS DOMÍNIOS.....</b>	<b>25</b>
<b>AVALIAÇÃO DA ADEQUAÇÃO DO TAMANHO DA AMOSTRA À POF .....</b>	<b>28</b>
<b>SELEÇÃO DA AMOSTRA.....</b>	<b>31</b>
DEFINIÇÃO DOS GRUPOS DE ROTAÇÃO.....	31
PROCEDIMENTOS OPERACIONAIS DE SELEÇÃO.....	32
<b>ROTAÇÃO DA AMOSTRA.....</b>	<b>33</b>
ROTAÇÃO DA AMOSTRA DE SETORES .....	33
ATUALIZAÇÃO DO CADASTRO MESTRE.....	34
<b>CONCLUSÕES .....</b>	<b>35</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>35</b>
<b>ANEXO 1 - ESTRATIFICAÇÃO GEOGRÁFICA: TABELAS E MAPAS.....</b>	<b>38</b>
<b>ANEXO 2 – ESQUEMA DE ROTAÇÃO DE DOMICÍLIOS.....</b>	<b>65</b>
<b>ANEXO 3 - TAMANHO DA AMOSTRA MESTRA POR DOMÍNIOS.....</b>	<b>66</b>



# Apresentação

O IBGE encontra-se no momento em fase de desenvolvimento de um projeto de reformulação de suas principais pesquisas domiciliares por amostragem probabilística. São destaques desta reformulação a integração da PNAD - Pesquisa Nacional por Amostra de Domicílios e da PME - Pesquisa Mensal de Emprego numa única pesquisa domiciliar contínua e a construção de uma amostra única para ser usada como base para esta e para as demais pesquisas domiciliares realizadas pelo instituto, tais como a POF - Pesquisa de Orçamentos Familiares e ECINF - Pesquisa de Economia Informal Urbana.

Os planos amostrais dessas pesquisas possuem várias semelhanças, como por exemplo: o uso de amostragem de conglomerados em mais de um estágio de seleção, passando primeiramente pela seleção de setores censitários e posteriormente a dos domicílios; a adoção de algum tipo de estratificação dos setores censitários; a seleção de setores com probabilidade proporcional ao tamanho, para dar conta da grande variabilidade do tamanho dos setores em número de domicílios; a utilização do mesmo cadastro de setores, a base territorial dividida em setores censitários; a realização da operação de atualização da listagem de domicílios nos setores selecionados para a amostra, antes da seleção dos domicílios para compor a amostra de cada pesquisa.

Tirar proveito dessas semelhanças é a motivação principal para a definição de um cadastro mestre e a seleção de uma amostra única, chamada Amostra Mestre, para atender a todas as pesquisas que já fazem parte do programa regular de pesquisas domiciliares do IBGE, além de possibilitar o atendimento de demandas específicas de forma integrada em um sistema de pesquisas domiciliares.

Este documento contém a descrição dos procedimentos usados para a definição do plano amostral dessa Amostra Mestre e da sua primeira utilização, qual seja, a seleção da amostra da Pesquisa de Orçamentos Familiares, prevista para ir a campo no primeiro trimestre de 2008.

**Sonia Albieri**  
Coordenação de Métodos e Qualidade



# Introdução

Na proposta de um novo sistema de pesquisas domiciliares por amostragem no IBGE está sendo discutida a integração das pesquisas. Essa integração aplica-se aos conceitos e definições de variáveis comuns, aos procedimentos e equipe de coleta, aos sistemas e métodos de apuração, crítica e imputação dos dados e, de forma inédita no Brasil, baseia-se no uso de um mesmo cadastro de seleção e de uma amostra comum.

Este texto trata do último aspecto de integração listado, a definição de uma amostra comum, denominada Amostra Mestra. Esta amostra corresponde a um conjunto de unidades de área selecionadas de um cadastro, segundo um método probabilístico de seleção, a partir da qual seja possível selecionar subamostras para atender às diversas pesquisas.

A seleção das subamostras pode ser realizada adotando-se diferentes procedimentos. Para cada pesquisa do novo Sistema Integrado de Pesquisas Domiciliares (SIPD) é possível selecionar uma subamostra de unidades de área ou uma subamostra de domicílios em todas as unidades de área que compõem a Amostra Mestra. As subamostras podem ser selecionadas de forma independente ou com um certo controle para que tenham ou não algumas unidades coincidentes.

Dentre as principais pesquisas do SIPD, já definidas, está a Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) que integrará as atuais pesquisas PME (Pesquisa Mensal de Emprego) e PNAD (Pesquisa Nacional por Amostra de Domicílios) e terá abrangência nacional e periodicidade de divulgação trimestral. A Pesquisa de Orçamentos Familiares (POF), também com abrangência nacional e com previsão de ser realizada a cada cinco anos. A Pesquisa de Orçamentos Familiares Simplificada (POFs) planejada para ser realizada nos anos em que não houver POF e a Pesquisa de Economia Informal Urbana (ECINF) com periodicidade quinzenal.

Convém ressaltar que a origem desses estudos remonta ao ano de 2001 no qual Silva et al. apresenta em seu artigo intitulado “Algumas idéias para a revisão das pesquisas domiciliares por amostragem do IBGE” a primeira abordagem relativo ao tema. Outro trabalho desenvolvido anteriormente por Silva (1998), apontou questões referentes ao esquema de rotação da Pesquisa Mensal de Emprego. O resultado alcançado nesse estudo, de certa forma, foi incorporado no começo dos estudos para o planejamento da Amostra Mestra.

A seguir são descritos os aspectos de amostragem considerados no desenvolvimento e definição da Amostra Mestra: a população alvo, a abrangência geográfica, o plano amostral e o tamanho da amostra.

## **População alvo e abrangência geográfica**

A população alvo da Amostra Mestra inclui toda a população a ser investigada em todas as pesquisas. É constituída pelos moradores residentes em todos os domicílios na área que constitui a abrangência geográfica definida a seguir.

A abrangência geográfica da Amostra Mestra considera o âmbito das diversas pesquisas que farão uso dessa amostra comum. Assim, não se pode deixar de incluir qualquer parte do território que seja contemplado por alguma das pesquisas. Portanto, a abrangência geográfica da Amostra Mestra é constituída pelos setores censitários da Base Operacional Geográfica de 2000 compatibilizada com a malha municipal de 2001 de todo o território nacional.

Cabe ressaltar que, apesar de abranger todo o território nacional, algumas áreas foram excluídas por questões operacionais. São áreas com pequeno número de domicílios e de pessoas e com características especiais. Além de não fazerem parte do âmbito das atuais pesquisas domiciliares do IBGE. Foram excluídas as áreas definidas pelo IBGE como sendo quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos e hospitais.

## **Plano amostral**

A Amostra Mestra, como citado anteriormente, é um conjunto de unidades de áreas que são selecionadas para atender a diversas pesquisas. Essas unidades são consideradas unidades primárias de amostragem (UPAs) no planejamento amostral de cada uma das pesquisas que utilizarão a Amostra Mestra.

O plano amostral escolhido foi um desenho estratificado, com seleção das unidades primárias com probabilidade proporcional a uma medida de tamanho. A opção por este plano se justifica pelo fato das metodologias de todas as pesquisas domiciliares do IBGE incluírem planos amostrais que empregam algum tipo de estratificação das UPAs e seleção destas unidades com probabilidade proporcional ao tamanho, conforme descrito em Bianchini e Albieri (2003).

## **Cadastro de seleção**

O cadastro para seleção da Amostra Mestra é um cadastro denominado Cadastro Mestre, que corresponde a um conjunto de unidades de áreas com seus limites bem definidos, que cobrem todo o país. Para cada uma dessas unidades são associadas informações sobre divisões administrativas, contagens populacionais e outras características sociodemográficas.

Como unidades básicas do cadastro foram escolhidos os setores censitários, pois são as menores partições do território para as quais são conhecidas as informações necessárias para sua construção.

O sistema de referência (cadastro) utilizado foi um arquivo contendo informações provenientes do Censo Demográfico 2000 sobre os setores censitários da abrangência geográfica, cujos limites são definidos na Base Operacional Geográfica de 2000 compatibilizada com a malha municipal de 2001, totalizando 214.836 setores (o número de setores nas áreas excluídas é de 983).

### **Definição das unidades primárias de amostragem**

Um importante aspecto para seleção de uma Amostra Mestra é a definição de suas unidades primárias de amostragem (UPAs). Em pesquisas domiciliares tais unidades são definidas, em geral, por unidades de área com um determinado tamanho mínimo populacional contabilizado em termos de domicílios ou pessoas.

As UPAs podem ser definidas como sendo as unidades básicas do Cadastro Mestre, ou podem corresponder a agregações contíguas destas, como por exemplo uma divisão administrativa. Não é interessante definir a UPA como uma unidade muito pequena, pois esta pode não possuir um tamanho suficiente (de domicílios ou pessoas) para o uso que se fará da amostra. Também não é recomendável definir a UPA como sendo uma unidade muito grande, já que o número de UPAs pode ficar reduzido, o que não justificaria a seleção de uma amostra.

Para a Amostra Mestra desenhada para o Sistema Integrado de Pesquisas Domiciliares do IBGE foi definida como UPA a unidade básica do cadastro, ou seja, o setor censitário. O número de setores censitários no país, dentre os quais será selecionada a Amostra Mestra, é grande e seus tamanhos são, em média, suficientes para a utilização em diversas pesquisas. Outra motivação para a escolha do setor censitário como UPA foi o desejo de se obter um maior espalhamento espacial da amostra.

### **Estratificação**

A estratificação das UPAs, ou seja, dos setores censitários, da Amostra Mestra foi definida levando-se em consideração os objetivos das diversas pesquisas que serão contempladas por esta amostra e também as questões operacionais e domínios de divulgação.

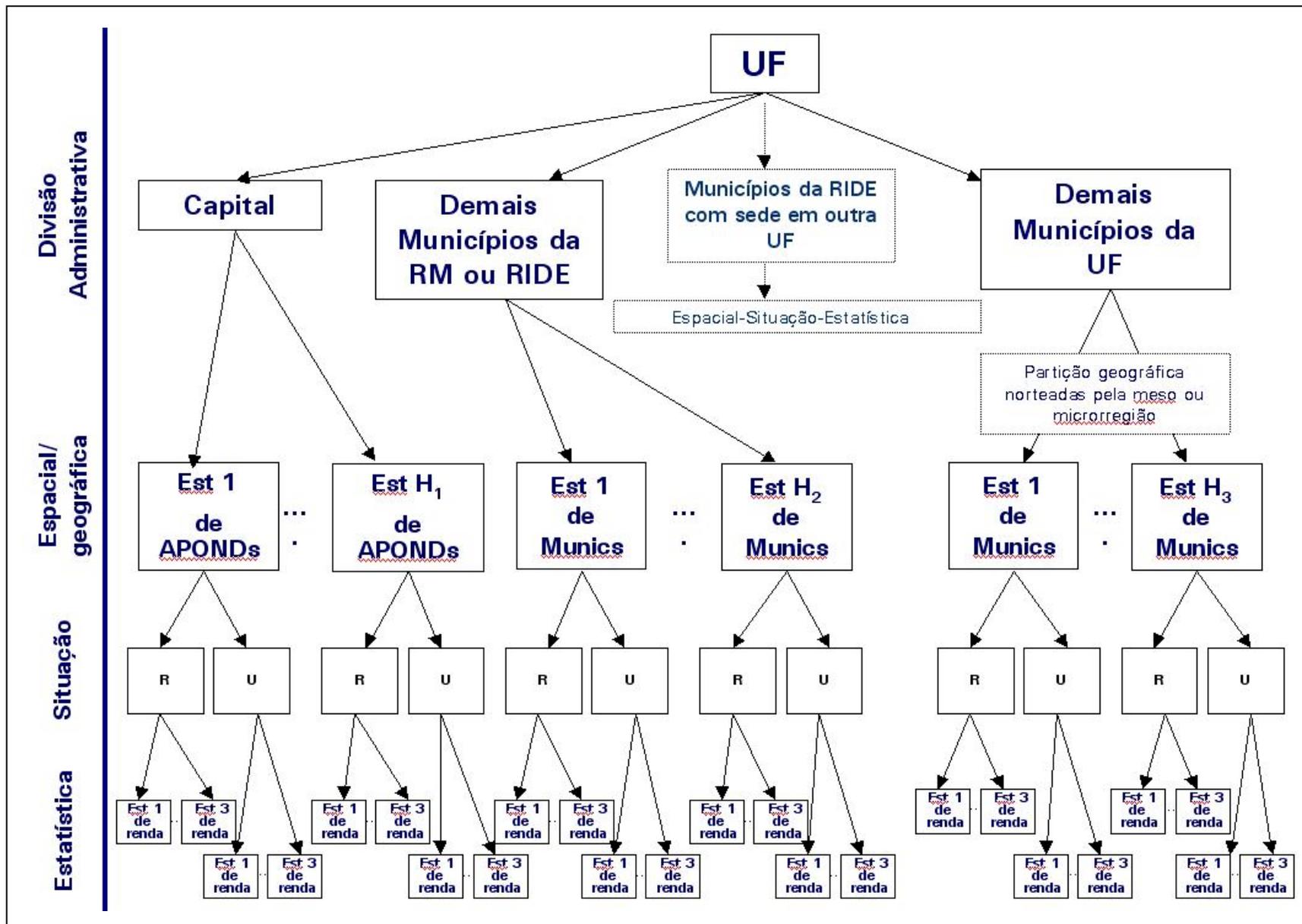
As UPAs foram estratificadas inicialmente por divisão administrativa, isto é associou-se a todos os setores de um município o seu respectivo estrato administrativo. Posteriormente, dentro de cada estrato definido na etapa anterior, foram criados estratos geográficos e espaciais de setores de acordo com um conjunto de características. Após esta etapa, os setores foram classificados diretamente em estratos considerando a característica

de pertencerem à parte urbana ou rural e em seguida, de acordo com a renda do responsável pelo domicílio, o que corresponde, respectivamente, à estratificação por situação e estratificação estatística do setor. Para maiores detalhes ver Freitas e Lila (2004).

Nas capitais, onde foi possível, considerando o seu tamanho em número de setores, foram criados estratos de áreas de ponderação do Censo 2000 (Apond), que são áreas contínuas definidas como a menor partição do território com resultados divulgados das características investigadas na amostra do censo. Somente após esta estratificação, foi aplicado o mesmo procedimento para estratificar os setores tal como definido para os demais estratos de municípios.

A figura 1, a seguir, ilustra o esquema de estratificação e, nas próximas seções, as etapas de estratificação são descritas com mais detalhes.

Figura 1 - Esquema de estratificação



## **Tamanho mínimo dos estratos**

Para definição dos limites dos estratos foi preciso decidir o tamanho mínimo, tanto para os estratos geográficos e espaciais, quanto para os estratos estatísticos. Essa definição foi importante para evitar a formação de estratos muito pequenos, ocasionando frações de amostragem muito grandes. Assim, a primeira decisão foi considerar como medida de tamanho o número de setores.

Na última etapa de estratificação, o tamanho mínimo dos estratos foi definido considerando o esquema de rotação da amostra de domicílios da Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC). De acordo com este esquema, é necessário construir uma quantidade de grupos mutuamente exclusivos de setores censitários, para atender a distribuição da amostra da PNADC ao longo do tempo.

Como é desejável que a amostra seja balanceada ao longo do tempo, isto é, que todos os grupos contenham pelo menos um setor da amostra de cada estrato final, o tamanho mínimo da amostra em cada estrato deve coincidir com a quantidade de grupos. Com isso, o tamanho mínimo do estrato deve ser tal que possibilite selecionar essa amostra de setores com uma fração amostral pequena (de no máximo 10%). O esquema de rotação definido requer 15 grupos, dessa forma foi escolhido como tamanho mínimo 150 setores por estrato final (veja seção *Esquema de rotação da amostra de domicílios na PNADC*).

Definido o tamanho mínimo dos estratos finais (estatísticos), foi preciso definir o tamanho mínimo dos estratos em cada situação (urbano e rural) de tal forma a permitir sua divisão em pelo menos 2 estratos estatísticos. O limite inferior do número de setores nos estratos foi fixado em 300. Esse limite variou por UF devido ao fato de algumas serem muito grandes em número de setores, o que ocasionaria a criação de um número excessivo de estratos, particionando demais o seu território. Ressalta-se que esses limites, em alguns casos, não foram respeitados, para evitar a definição de estratos que contivessem áreas com características muito diferentes, sendo, então, utilizados limites um pouco diferentes dos definidos.

No quadro 1 são apresentados os limites inferiores dos tamanhos dos estratos geográficos e espaciais por Unidade da Federação, os limites definidos no estudo e os efetivamente utilizados.

Quadro 1 - Limite inferior do tamanho dos estratos geográficos e espaciais em número de setores segundo a UF

UF	Limite inferior do tamanho dos estratos em número de setores	
	Definido	Efetivo
11 – RO	366	275
12 – AC	307	307
13 – AM	417	443
14 – RR	205	205
15 – PA	1 104	1 018
16 – AP	215	148
17 – TO	309	466
21 – MA	1 083	1 340
22 – PI	601	541
23 – CE	497	924
24 – RN	347	330
25 – PB	713	838
26 – PE	1 147	1 405
27 – AL	352	393
28 – SE	343	801
29 – BA	2 435	614
31 – MG	2 623	2 583
32 – ES	373	591
33 – RJ	1 502	1 268
35 – SP	4 123	2 344
41 – PR	1 433	1 834
42 – SC	892	1 113
43 – RS	2 284	3 486
50 – MS	399	455
51 – MT	416	311
52 – GO	657	634
53 – DF	381	381

### Estratificação por divisão administrativa

Um dos motivos para utilizar estratificação no planejamento amostral de uma pesquisa é que, às vezes, razões administrativas ou até mesmo legais geram demandas por estimativas para determinadas regiões geográficas e podem, também, impor que estas regiões sejam especificamente contempladas no plano amostral. Ao considerar cada divisão destas como um estrato, garante-se que nelas serão selecionadas unidades amostrais, com possibilidade de controlar a precisão das estimativas.

As principais pesquisas do novo sistema integrado estão sendo planejadas de modo a fornecer resultados para cada uma das Unidades da Federação (UFs), o que torna necessário garantir a seleção de uma amostra em cada uma delas. Por isso, num primeiro passo, determinou-se que cada uma seria tratada como um estrato.

Dentro de cada UF, os municípios foram também classificados em estratos. Tal procedimento visa aprimorar a precisão, bem como garantir que certas divisões administrativas da UF fossem contempladas na amostra. Como não se sabe, a priori, se haverá a necessidade de divulgação de resultados para estas divisões, decidiu-se considerá-las nesta etapa de estratificação. As divisões administrativas consideradas foram as Regiões Metropolitanas<sup>1</sup> (RM) e as Regiões Integradas de Desenvolvimento (RIDE) de Teresina e do Distrito Federal.

Nesta etapa de estratificação, os municípios foram classificados em três grupos:

- (1) Capital;
- (2) Demais municípios pertencentes à RM ou à RIDE e
- (3) Demais municípios da UF.

Nas UFs do Maranhão, de Minas Gerais e de Goiás foi definido um quarto grupo que contém os municípios que pertencem à RIDE de Teresina ou do Distrito Federal.

### **Estratificação geográfica e espacial**

Após a classificação dos municípios considerando algumas divisões administrativas, os municípios do grupo (3) (demais municípios da UF) foram classificados em estratos geográficos. O pequeno número de estratos em cada UF, devido à restrição do tamanho mínimo, não permitiu maior detalhamento geográfico. Procurou-se apenas garantir que a delimitação não separasse áreas muito semelhantes ou, ao contrário, juntasse áreas muito diferentes, tomando como base as divisões em meso e microrregiões. Nas áreas em que ocorreram mudanças significativas na organização do território depois de 1989 (o ano de estabelecimento das meso e microrregiões), foram utilizados conhecimentos derivados de outras pesquisas da Coordenação de Geografia, da Diretoria de Geociências.

No grupo (1) (capital) foi aplicado um método de estratificação espacial para classificar as áreas de ponderação (Aponds) do Censo 2000 e no grupo (2) (demais municípios da RM ou RIDE) foi aplicado o mesmo método para classificar os municípios. Cabe ressaltar que as Aponds foram definidas pelas administrações municipais ou por critérios de homogeneidade estabelecidos pelo próprio IBGE.

As características utilizadas no grupo (1) foram a renda média dos responsáveis pelos domicílios e a taxa de desocupação, e no grupo (2), além destas

---

<sup>1</sup> Foram apenas consideradas as Regiões Metropolitanas que contêm o município da capital.

duas características, foram consideradas também a densidade demográfica e a proporção de ocupados em atividades agrícolas. Para maiores detalhes ver Freitas e Lila (2004).

O método usado para definir esta estratificação espacial é inédito no IBGE. Consiste em classificar as unidades em grupos que sejam, ao mesmo tempo, homogêneos segundo as características escolhidas e contíguos. Ou seja, os estratos são formados por unidades vizinhas e "parecidas" (ver Freitas, 2002). A aplicação deste método foi realizada utilizando o pacote SKATER conforme descrito em Assunção (2000), sendo que o tamanho mínimo considerado na definição de cada estrato foi o limite da UF correspondente, já citado no quadro 1.

Considerar os limites geográficos no procedimento de estratificação pode contribuir para que as pesquisas forneçam informações úteis para diversas prefeituras.

Um outro aspecto relacionado com este tipo de estratificação espacial é que ele pode facilitar a operação de coleta, principalmente em pesquisas contínuas realizadas com determinada periodicidade, pois permite a alocação de equipes permanentes em cada uma destas regiões, agilizando e possibilitando maior controle das atividades da operação de campo, além de reduzir custos de deslocamento entre as unidades da amostra.

A tabela 1 a seguir descreve os números de estratos geográficos e espaciais por UF e a tabela A1 no anexo 1 contém a distribuição do número de setores, domicílios e pessoas por estrato. Adicionalmente, para ilustrar esta estratificação, são apresentados, no anexo 1, os mapas de todas as UFs com a respectiva definição dos estratos.

Tabela 1 - Número de estratos geográficos e espaciais por tipo segundo a UF

(continua)

UF	Número de estratos geográficos e espaciais					
	Total	Capital	Demais Municípios da RM ou RIDE	Colar Metropolitano	Municípios da RIDE com sede em outra UF	Demais Municípios da UF
<b>Brasil</b>	<b>144</b>	<b>40</b>	<b>21</b>	<b>1</b>	<b>4</b>	<b>78</b>
11 – RO	3	1	0	0	0	2
12 – AC	2	1	0	0	0	1
13 – AM	5	2	0	0	0	3
14 – RR	2	1	0	0	0	1
15 – PA	4	1	1	0	0	2
16 – AP	3	1	1	0	0	1
17 – TO	4	1	1	0	0	2

Tabela 1 - Número de estratos geográficos e espaciais por tipo segundo a UF

(conclusão)

UF	Número de estratos geográficos e espaciais					
	Total	Capital	Demais Municípios da RM ou RIDE	Colar Metropolitano	Municípios da RIDE com sede em outra UF	Demais Municípios da UF
21 – MA	6	1	1	0	1	3
22 – PI	5	1	0	0	1	3
23 – CE	8	3	1	0	0	4
24 – RN	6	1	1	0	0	4
25 – PB	4	1	1	0	0	2
26 – PE	5	1	1	0	0	3
27 – AL	5	1	1	0	0	3
28 – SE	3	1	0	0	0	2
29 – BA	8	2	1	0	0	5
31 – MG	10	2	1	1	1	5
32 – ES	5	1	1	0	0	3
33 – RJ	10	3	3	0	0	4
35 – SP	10	3	2	0	0	5
41 – PR	6	2	1	0	0	3
42 – SC	5	1	1	0	0	3
43 – RS	6	2	1	0	0	3
50 – MS	4	1	0	0	0	3
51 – MT	6	1	0	0	0	5
52 – GO	6	1	1	0	1	3
53 – DF	3	3	0	0	0	0

### Estratificação por situação do setor

Após as etapas iniciais de estratificação, administrativa, geográfica e espacial, quando foram classificados os municípios e as Aponds, associou-se a cada setor o estrato do município ou da Apond a que pertence. Após essa associação, em cada estrato definido anteriormente, os setores foram estratificados diretamente, sendo classificados em grupos, primeiro considerando sua situação.

Essa etapa de estratificação foi elaborada para garantir a seleção de setores censitários tanto na área urbana, quanto na área rural, em cada UF, pois algumas das pesquisas que utilizarão a Amostra Mestra têm como objetivo estimar características separadamente para cada uma destas áreas. Para isso, os setores em cada estrato

geográfico e espacial foram classificados em 2 grupos de acordo com sua situação, sendo criados, portanto, um estrato urbano e um estrato rural. O número de estratos por situação em cada UF são apresentados na tabela 2.

Note que o número de estratos urbanos corresponde ao número de estratos geográficos e espaciais, já o número de estratos rurais é um pouco inferior. Esse fato indica que em todos os estratos havia pelo menos 1 setor urbano, podendo ser criado o correspondente estrato urbano, e que em 20 estratos não havia nenhum setor rural, não sendo possível a criação do correspondente estrato rural.

Tabela 2 - Número de estratos por situação segundo a UF

UF	Número de estratos	
	Urbano	Rural
<b>Brasil</b>	<b>144</b>	<b>124</b>
11 – RO	3	3
12 – AC	2	2
13 – AM	5	4
14 – RR	2	2
15 – PA	4	4
16 – AP	3	3
17 – TO	4	4
21 – MA	6	6
22 – PI	5	5
23 – CE	8	5
24 – RN	6	5
25 – PB	4	3
26 – PE	5	4
27 – AL	5	5
28 – SE	3	2
29 – BA	8	7
31 – MG	10	8
32 – ES	5	4
33 – RJ	10	6
35 – SP	10	10
41 – PR	6	4
42 – SC	5	5
43 – RS	6	5
50 – MS	4	4
51 – MT	6	6
52 – GO	6	6
53 – DF	3	2

## Estratificação estatística

A última etapa de estratificação foi a estatística, na qual os setores censitários foram classificados em grupos homogêneos segundo algumas de suas características conhecidas para todas as unidades da população sob estudo. As características consideradas foram a renda total dos responsáveis pelos domicílios e a taxa de desocupação. Estas duas características foram tratadas separadamente e em conjunto, isto é, de forma univariada e simultânea.

Optou-se pelo uso da renda, pois acredita-se que a taxa de desocupação tem um comportamento mais instável ao longo da década, além de, reconhecidamente, a renda ter relação direta com as características de interesse investigadas na maioria das pesquisas domiciliares. Uma justificativa para o emprego desta estratificação é que se pode aumentar a precisão das estimativas para o conjunto da população.

O método utilizado consiste em classificar os setores em grupos de tal forma que minimize a variância do estimador do total da característica de interesse, considerando o plano amostral comumente utilizado nas pesquisas domiciliares: amostragem conglomerada com seleção das UPAs (setores) com probabilidade proporcional a uma medida de tamanho (número de domicílios).

A medida de distância associada a pares de setores utilizada no método é dada por

$$d(i, i') = N_i \cdot N_{i'} \cdot \left( \frac{Y_i}{N_i} - \frac{Y_{i'}}{N_{i'}} \right)^2$$

onde

$N_i$  é o número de domicílios no setor  $i$ ;

$N_{i'}$  é o número de domicílios no setor  $i'$ ;

$Y_i$  é total da característica de interesse  $y$  no setor  $i$ ;

$Y_{i'}$  é total da característica de interesse  $y$  no setor  $i'$  e

$y$  é a renda dos responsáveis pelos domicílios.

O número de estratos estatísticos formados em cada estrato de situação foi de 2 ou de 3 estratos e 150 setores foi o tamanho mínimo dos estratos. Cabe ressaltar que alguns estratos de situação não foram divididos em estratos estatísticos por não possuírem tamanho suficiente para a aplicação deste procedimento. A implementação deste método foi feita utilizando algoritmos de otimização desenvolvidos por Montenegro e Brito (2006).

O número de estratos estatísticos definidos em cada Unidade da Federação é apresentado na tabela 3.

Tabela 3 - Número de estratos por situação e estatísticos segundo a UF

UF	Número de estratos	
	Por Situação	Estatísticos
<b>Brasil</b>	<b>268</b>	<b>550</b>
11 – RO	6	11
12 – AC	4	4
13 – AM	9	13
14 – RR	4	4
15 – PA	8	14
16 – AP	6	6
17 – TO	8	10
21 – MA	12	24
22 – PI	10	19
23 – CE	13	36
24 – RN	11	13
25 – PB	7	16
26 – PE	9	25
27 – AL	10	13
28 – SE	5	9
29 – BA	15	36
31 – MG	18	45
32 – ES	9	14
33 – RJ	16	37
35 – SP	20	51
41 – PR	10	29
42 – SC	10	23
43 – RS	11	30
50 – MS	8	13
51 – MT	12	18
52 – GO	12	28
53 – DF	5	9

## Tamanho da amostra

O tamanho da amostra foi determinado considerando a precisão desejada para a estimativa de um parâmetro de interesse. O parâmetro escolhido para este estudo foi o “total de pessoas desocupadas no trimestre”, que será um dos principais indicadores da PNADC.

Este total pode ser definido como

$$Y = \sum_{h=1}^L \sum_{i=1}^{M_h} \sum_{j=1}^{N_{hi}} y_{hij} = \sum_{h=1}^L Y_h, \quad (1)$$

onde

$h$  é o índice do estrato a que pertence o setor;

$i$  é o índice do setor dentro do estrato;

$j$  é o índice do domicílio dentro do setor;

$L$  é o número total de estratos;

$M_h$  é o número de setores na população do estrato  $h$  ;

$N_{hi}$  é o número de domicílios na população do setor  $i$  , do estrato  $h$  ;

$y_{hij}$  é o valor da variável  $y$  no domicílio  $j$  do setor  $i$  , do estrato  $h$  e

$Y_h$  é o total da variável  $y$  no estrato  $h$  .

Para definir um estimador para o parâmetro de interesse (1) e fazer os cálculos do tamanho de amostra foi preciso considerar o desenho amostral das pesquisas que farão uso da Amostra Mestra, especialmente a PNADC, que será um dos pilares do sistema. O desenho proposto para esta pesquisa é um desenho conglomerado em dois estágios, com estratificação das UPAs (setores censitários), seleção das UPAs com probabilidade proporcional ao tamanho e seleção das unidades secundárias de amostragem (USAs), os domicílios, com probabilidade igual e número fixo de unidades. Com base neste desenho amostral foi definido um estimador e sua respectiva variância através das expressões. Ver Kish (1965).

$$\hat{Y} = \sum_{h=1}^L \frac{1}{m_h} \sum_{i=1}^{m_h} \frac{\hat{Y}_{hi}}{p_{hi}} \quad (2)$$

e

$$V(\hat{Y}) = \sum_{h=1}^L \frac{1}{m_h} \cdot \left[ \sum_{i=1}^{M_h} \frac{Y_{hi}^2}{p_{hi}} - Y_h^2 + \sum_{i=1}^{M_h} \frac{N_{hi}^2}{p_{hi}} \cdot \frac{S_{yhi}^2}{n_{hi}} \cdot \frac{N_{hi} - n_{hi}}{N_{hi}} \right], \quad (3)$$

onde

$m_h$  é o número de setores na amostra do estrato  $h$  ,

$p_{hi} = \frac{N_{hi}}{N_h}$  é o tamanho relativo do setor  $i$ , do estrato  $h$ , que define a probabilidade de seleção deste setor em um sorteio, caso a amostra de setores fosse selecionada com PPT com reposição;

$N_h$  é o número de domicílios na população do estrato  $h$ ;

$\hat{Y}_{hi} = \sum_{j=1}^{n_{hi}} w_{j|hi} \cdot y_{hij}$  é o estimador simples do total  $Y_{hi} = \sum_{j=1}^{N_{hi}} y_{hij}$  da variável  $y$  no setor  $i$  do estrato  $h$ ;

$n_{hi}$  é o número de domicílios na amostra do setor  $i$ , do estrato  $h$ ;

$w_{j|hi} = \frac{N_{hi}}{n_{hi}}$  é o peso do domicílio  $j$  dado a seleção do setor  $i$  do estrato  $h$  e

$S_{yhi}^2 = \frac{1}{N_{hi} - 1} \cdot \sum_{j=1}^{N_{hi}} (y_{hij} - \bar{Y}_{hi})^2$  é a variância da variável de interesse  $y$  na população no setor  $i$  do estrato  $h$ .

Como a característica de interesse só foi investigada no Censo 2000 através do questionário da amostra, não foi possível utilizar a fórmula da variância deste estimador descrita na expressão (3) para calcular o tamanho da amostra de setores e de domicílios. Foi preciso definir um estimador não viciado para esta variância com base nos dados da amostra do Censo 2000, conforme Lima e Bianchini (1986), de modo a dimensionar a amostra necessária para estimar o parâmetro de interesse com a precisão desejada.

O estimador utilizado é dado por:

$$\hat{V}(\hat{Y}) = \sum_{h=1}^L \frac{1}{m_h} \cdot \left[ \sum_{i=1}^{M_h} \frac{\tilde{Y}_{hi}^2}{p_{hi}} - \tilde{Y}_h^2 + \sum_{i=1}^{M_h} \frac{N_{hi}^2}{p_{hi}} \cdot \frac{s_{yhi}^2}{n_{hi}} \cdot \frac{N_{hi} - n_{hi}}{N_{hi}} \right], \quad (4)$$

onde

$$\tilde{Y}_{hi}^2 = N_{hi}^2 \cdot \left( \bar{y}_{hi}^2 - \frac{N_{hi} - n_{hi}^*}{N_{hi}} \cdot \frac{s_{yhi}^2}{n_{hi}^*} \right) \text{ é um estimador não viciado para } Y_{hi}^2;$$

$$\tilde{Y}_h^2 = \left( \sum_{i=1}^{M_h} N_{hi} \cdot \bar{y}_{hi} \right)^2 - \sum_{i=1}^{M_h} \left( N_{hi}^2 \cdot \frac{N_{hi} - n_{hi}^*}{N_{hi}} \cdot \frac{s_{yhi}^2}{n_{hi}^*} \right) \text{ é um estimador não viciado para } Y_h^2;$$

$$s_{yhi}^2 = \frac{1}{n_{hi}^* - 1} \cdot \sum_{j=1}^{n_{hi}^*} (y_{hij} - \bar{y}_{hi})^2 \text{ e}$$

$n_{hi}^*$  é o número de domicílios na amostra do setor  $i$ , do estrato  $h$  no Censo 2000.

Considerando como fixo o número de domicílios a serem selecionados por setor ( $\bar{n}$ ) e com base no estimador da variância descrito anteriormente em (4), calcula-se o tamanho da amostra total de setores por meio da seguinte expressão

$$m = \frac{N \cdot \sum_{h=1}^H \sum_{i=1}^{M_h} \left\{ N_{hi} \cdot (\bar{y}_{hi}^2 - \bar{y}_h^2) - \left[ \left( \frac{N_{hi} - n_{hi}^*}{n_{hi}^*} \right) \cdot \left( 1 - \frac{N_{hi}}{N_h} \right) - \left( \frac{N_{hi} - \bar{n}}{\bar{n}} \right) \right] \cdot s_{yhi}^2 \right\}}{(CV(\hat{Y}) \cdot \hat{Y})^2}, \text{ onde}$$

$CV(\hat{Y})$  é o coeficiente de variação desejado para a estimativa de interesse e

$\hat{Y}$  é a estimativa do total da característica de interesse.

Para calcular efetivamente o tamanho da amostra de setores foi preciso escolher qual a precisão desejada para a estimativa do parâmetro de interesse, para qual domínio esta estimativa seria produzida e com qual periodicidade, e o número de domicílios a serem selecionados por setor.

Na PNADC deverão ser divulgados resultados para cada uma das Unidades da Federação separadamente, por isso o estudo considerou que a estimativa do total de pessoas desocupadas será calculada para cada UF, e com periodicidade trimestral, pois a coleta dos dados será feita em um ciclo de 3 meses. Esse período foi definido por questões operacionais e de custo, já que seria impraticável coletar a amostra inteira todo mês. Além disso, os resultados obtidos por Freitas e Lila (2003) nos estudos sobre a significância estatística das diferenças dos indicadores mensais produzidos pela atual pesquisa sobre mercado de trabalho do IBGE indicam que apenas 13% das diferenças foram consideradas significativas, dentre estas, 33% envolvem os meses de dezembro e janeiro, meses fortemente influenciados por questões sazonais.

O nível de precisão, expresso em termos do CV, foi determinado diferentemente para cada UF, levando em consideração os níveis de precisão obtidos com a atual amostra da PNAD. O número de domicílios a serem selecionados por setor foi fixado em 14, pois em estudos realizados, foi observado que o aumento neste número não acarretaria uma redução considerável no tamanho da amostra de setores, pelo fato de haver grande homogeneidade entre os domicílios de um mesmo setor (ver Freitas e Lila, 2004).

## Esquema de rotação da amostra de domicílios na PNADC

Para calcular o tamanho da amostra foi necessário definir o esquema de rotação da amostra de domicílios que será adotado na PNADC, pois este esquema define o tamanho mínimo de amostra em cada estrato.

A PNADC está sendo planejada para ter periodicidade de coleta trimestral, ou seja, a amostra total de setores e domicílios será coletada em um período de 3 meses, para ao final deste ciclo serem produzidas as estimativas dos indicadores de interesse. Sendo assim, um possível esquema de rotação da amostra a ser adotado pode ser uma variação do esquema 1-2(v), onde v seria o número de visitas a serem realizadas. Neste esquema o domicílio é entrevistado 1 mês e sai da amostra por 2 meses seguidos, sendo esta seqüência repetida v vezes.

O esquema de rotação proposto foi comparado com outros esquema por Yansaneh (1998), mostrando ser mais eficiente no caso em que o principal interesse é a inferência a respeito de mudanças no comportamento de indicadores, considerando períodos de três meses. O número de visitas definido para a pesquisa foi de 5 visitas, portanto o esquema a ser adotado será o 1-2(5), cuja ilustração encontra-se no anexo 2.

Ressalta-se que a decisão de adotar o esquema 1-2(5) foi tomada após ampla consulta a usuários dos dados da pesquisa, que evidenciaram a necessidade de poder comparar as informações de uma mesma pessoa em anos consecutivos. Esse esquema propicia essa comparação, pois há a sobreposição de 25% da amostra de um trimestre para o mesmo trimestre do ano seguinte (sobreposição esperada).

## Alocação do tamanho da amostra

Calculado o tamanho total da amostra de setores, foi preciso alocá-lo nos estratos. A alocação foi feita proporcionalmente ao número de domicílios, através da expressão

$$m_h = m \cdot \frac{N_h}{N}.$$

Devido às restrições de tamanho mínimo da amostra em cada estrato, foi preciso fazer alguns ajustes nos tamanhos de forma a garantir que o número de setores em cada estrato fosse no mínimo igual ao número de grupos necessário para o esquema de rotação escolhido para a amostra de domicílios da PNADC. Com isso, o tamanho da amostra em cada estrato passou a ser definido por

$$m_h^* = \max\left(m \cdot \frac{N_h}{N}, 15\right).$$

Após esses ajustes em cada estrato estatístico (final), outro ajuste também foi necessário no nível do estrato geográfico e espacial. Como a pesquisa será contínua, um dos benefícios esperados é a homogeneidade da carga de trabalho ao longo do tempo,

evitando sobrecarga em apenas alguns meses do ano, como ocorre atualmente. Para isso seria interessante que o número de setores na amostra em cada área da pesquisa fosse igual em todos os períodos de coleta de dados.

Em uma pesquisa com periodicidade de coleta trimestral pode-se dividir o tamanho da amostra igualmente pelos meses ou semanas do trimestre, para que em cada um destes períodos a carga de trabalho seja a mesma. Desta forma, o tamanho da amostra em cada estrato geográfico e espacial foi ajustado para ser múltiplo de 3, considerando-se uma distribuição igual nos meses.

Os tamanhos da amostra estão na tabela 4, assim como o nível de precisão esperado da estimativa do total de pessoas desocupadas para cada Unidade da Federação e por situação do setor.

Tabela 4 - Tamanho da Amostra Mestra para estimar o total de desocupados e o respectivo coeficiente de variação esperado por situação de setores segundo o Total Brasil, Grandes regiões e Unidades da Federação por ciclo de acumulação trimestral. (continua)

Brasil / Grande Região / UF	Tamanho da amostra de setores					
	Total	CV (%)	Situação dos Setores			
			Urbano		Rural	
			Total	CV (%)	Total	CV (%)
<b>BRASIL</b>	<b>12 800</b>	<b>1,0</b>	<b>8 764</b>	<b>1,1</b>	<b>4 036</b>	<b>3,0</b>
<b>NORTE</b>	<b>1 296</b>	<b>2,7</b>	<b>864</b>	<b>2,8</b>	<b>432</b>	<b>10,8</b>
11 – RO	216	6,7	119	7,0	97	20,1
12 – AC	105	8,8	73	8,7	32	32,7
13 – AM	285	4,7	203	4,8	82	24,5
14 – RR	81	8,4	58	8,7	23	29,0
15 – PA	336	5,0	217	5,1	119	17,5
16 – AP	69	8,3	55	8,4	14	48,1
17 – TO	204	6,7	139	7,0	65	20,7
<b>NORDESTE</b>	<b>4 838</b>	<b>1,5</b>	<b>3 135</b>	<b>1,5</b>	<b>1 703</b>	<b>5,0</b>
21 – MA	837	3,5	492	3,7	345	9,6
22 – PI	351	6,1	175	6,6	176	15,9
23 – CE	807	3,3	554	3,4	253	11,9
24 – RN	294	5,0	197	5,5	97	12,4
25 – PB	339	4,9	219	5,2	120	15,7
26 – PE	612	3,3	417	3,4	195	12,3
27 – AL	572	3,5	403	3,7	169	11,8
28 – SE	237	5,2	177	5,4	60	17,0
29 – BA	789	3,2	501	3,3	288	10,7

Nota: Os cálculos foram efetuados utilizando a base de dados do Censo Demográfico 2000.

Tabela 4 - Tamanho da Amostra Mestra para estimar o total de desocupados e o respectivo coeficiente de variação esperado por situação de setores segundo o Total Brasil, Grandes regiões e Unidades da Federação por ciclo de acumulação trimestral. (conclusão)

Brasil / Grande Região / UF	Tamanho da amostra de setores					
	Total	CV (%)	Situação dos Setores			
			Urbano		Rural	
			Total	CV (%)	Total	CV (%)
<b>SUDESTE</b>	<b>3 306</b>	<b>1,8</b>	<b>2 415</b>	<b>1,8</b>	<b>891</b>	<b>4,9</b>
31 – MG	993	2,9	680	3,0	313	12,1
32 – ES	453	3,5	362	3,5	91	17,6
33 – RJ	777	2,8	674	2,9	103	10,7
35 – SP	1083	2,8	699	2,9	384	5,2
<b>SUL</b>	<b>2 007</b>	<b>2,0</b>	<b>1 419</b>	<b>2,1</b>	<b>588</b>	<b>7,4</b>
41 – PR	675	3,3	466	3,4	209	11,8
42 – SC	633	3,5	467	3,6	166	12,8
43 – RS	699	3,4	486	3,5	213	13,2
<b>CENTRO-OESTE</b>	<b>1 353</b>	<b>2,4</b>	<b>931</b>	<b>2,5</b>	<b>422</b>	<b>8,3</b>
50 – MS	267	5,1	186	5,2	81	18,3
51 – MT	378	4,9	239	5,1	139	16,7
52 – GO	513	4,2	334	4,4	179	13,1
53 – DF	195	4,9	172	5,0	23	19,9

Nota: Os cálculos foram efetuados utilizando a base de dados do Censo Demográfico 2000.

## Avaliação da precisão da estimativa de outros parâmetros e para outros domínios

Após a determinação do tamanho da amostra de setores necessário para estimar o total de pessoas desocupadas no trimestre em cada UF, foi avaliada a precisão da estimativa deste mesmo parâmetro para outros domínios, Regiões Metropolitanas que contêm a capital, RIDEs e Município da capital. Também avaliou-se a precisão da estimativa do total da renda dos responsáveis pelos domicílios para todos os domínios em questão.

Após esta avaliação, foi feito um ajuste no tamanho da amostra do município de Palmas, pois não era suficiente para permitir estimar os parâmetros de interesse com um nível de precisão aceitável (coeficiente de variação de no máximo 15%).

Nas tabelas a seguir são apresentados o coeficiente de variação para a estimativa dos dois parâmetros de interesse e para os diversos domínios, e as tabelas do anexo 3 contêm os tamanhos de amostra para cada domínio.

Tabela 5 - Coeficiente de variação (%) estimado do estimador do total de desocupados e do estimador do total da renda do responsável por situação de setores segundo o Total Brasil, Grandes regiões e Unidades da Federação por ciclo de acumulação trimestral

Brasil / Grande Região / UF	Desocupados			Renda		
	Total	Urbano	Rural	Total	Urbano	Rural
<b>BRASIL</b>	<b>1,0</b>	<b>1,1</b>	<b>3,0</b>	<b>1,2</b>	<b>1,3</b>	<b>1,7</b>
<b>NORTE</b>	<b>2,7</b>	<b>2,8</b>	<b>10,8</b>	<b>3,6</b>	<b>4,2</b>	<b>5,7</b>
11 – RO	6,7	7,0	20,1	6,0	7,6	6,8
12 – AC	8,8	8,7	32,7	9,0	10,2	16,6
13 – AM	4,7	4,8	24,5	8,4	9,0	10,2
14 – RR	8,4	8,7	29,0	11,6	12,5	24,3
15 – PA	5,0	5,1	17,5	6,5	7,7	9,3
16 – AP	8,3	8,4	48,1	10,7	11,1	21,0
17 – TO	6,7	7,0	20,7	9,9	10,8	25,2
<b>NORDESTE</b>	<b>1,5</b>	<b>1,5</b>	<b>5,0</b>	<b>1,6</b>	<b>1,8</b>	<b>2,6</b>
21 – MA	3,5	3,7	9,6	4,9	6,0	4,9
22 – PI	6,1	6,6	15,9	6,7	7,9	4,3
23 – CE	3,3	3,4	11,9	3,8	4,2	3,9
24 – RN	5,0	5,5	12,4	5,9	6,3	15,2
25 – PB	4,9	5,2	15,7	5,0	5,6	5,0
26 – PE	3,3	3,4	12,3	3,7	4,0	6,4
27 – AL	3,5	3,7	11,8	7,6	8,6	7,1
28 – SE	5,2	5,4	17,0	7,0	7,7	8,6
29 – BA	3,2	3,3	10,7	3,5	3,9	6,8
<b>SUDESTE</b>	<b>1,8</b>	<b>1,8</b>	<b>4,9</b>	<b>2,0</b>	<b>2,1</b>	<b>3,4</b>
31 – MG	2,9	3,0	12,1	3,8	4,1	5,6
32 – ES	3,5	3,5	17,6	4,8	5,2	9,8
33 – RJ	2,8	2,9	10,7	3,2	3,3	11,0
35 – SP	2,8	2,9	5,2	3,0	3,1	5,2
<b>SUL</b>	<b>2,0</b>	<b>2,1</b>	<b>7,4</b>	<b>2,0</b>	<b>2,2</b>	<b>2,8</b>
41 – PR	3,3	3,4	11,8	3,7	4,0	5,2
42 – SC	3,5	3,6	12,8	3,3	3,6	5,9
43 – RS	3,4	3,5	13,2	3,1	3,5	4,0
<b>CENTRO-OESTE</b>	<b>2,4</b>	<b>2,5</b>	<b>8,3</b>	<b>3,0</b>	<b>3,2</b>	<b>7,8</b>
50 – MS	5,1	5,2	18,3	7,0	7,6	15,7
51 – MT	4,9	5,1	16,7	6,7	7,3	17,1
52 – GO	4,2	4,4	13,1	5,0	5,4	8,5
53 – DF	4,9	5,0	19,9	5,8	5,9	35,5

Nota: Os cálculos foram efetuados utilizando a base de dados do Censo Demográfico 2000.

Tabela 6 - Coeficiente de variação (%) estimado do estimador do total de desocupados e do estimador do total da renda dos responsáveis segundo as Regiões Metropolitanas e RIDEs de algumas UFs por ciclo de acumulação trimestral.

UF	Região Metropolitana / RIDE	Desocupados			Renda		
		Total	Urbano	Rural	Total	Urbano	Rural
15 – PA	Belém	6,2	6,3	29,0	9,9	10,0	23,4
16 – AP	Amapá	8,8	9,0	40,5	12,3	12,4	27,1
17 – TO	Palmas	9,9	10,0	68,2	16,7	17,3	36,8
21 – MA	São Luís	5,0	5,4	12,5	8,4	9,2	16,9
22 – RD	Teresina (RIDE)	7,8	8,0	30,0	8,0	8,3	18,0
23 – CE	Fortaleza	4,1	4,2	23,3	5,4	5,4	10,7
24 – RN	Natal	6,1	6,5	17,3	8,0	8,3	22,0
26 – PE	Recife	4,1	4,2	20,6	4,9	4,9	13,8
27 – AL	Maceió	4,3	4,3	37,4	11,6	11,7	26,7
29 – BA	Salvador	4,6	4,6	27,4	4,7	4,7	44,5
31 – MG	Belo Horizonte	4,5	4,5	20,5	6,9	6,9	37,3
32 – ES	Vitória	4,3	4,3	46,2	6,5	6,6	23,2
33 – RJ	Rio de Janeiro	3,3	3,3	30,0	3,7	3,7	44,2
35 – SP	São Paulo	3,8	4,0	6,0	4,3	4,4	5,9
41 – PR	Curitiba	5,2	5,4	16,9	4,6	4,7	14,2
42 – SC	Florianópolis	7,9	8,0	48,8	10,2	10,4	42,0
43 – RS	Porto Alegre	4,8	5,0	15,6	4,3	4,4	10,0
52 – GO	Goiânia	7,1	7,2	43,5	7,2	7,3	23,2
53 – DF	Distrito Federal (RIDE)	4,2	4,3	16,4	5,5	5,6	27,9

Nota: Os cálculos foram efetuados utilizando a base de dados do Censo Demográfico 2000.

Tabela 7 - Coeficiente de variação (%) estimado do estimador do total de desocupados e do estimador total da renda dos responsáveis por situação de setores segundo as capitais por ciclo de acumulação trimestral (continua)

UF	Capital	Desocupados			Renda		
		Total	Urbano	Rural	Total	Urbano	Rural
11 – RO	Porto Velho	9,9	11,2	19,1	10,7	11,6	14,1
12 – AC	Rio Branco	10,3	10,0	76,1	12,5	12,9	51,4
13 – AM	Manaus	5,4	5,4	55,8	10,6	10,6	46,1
14 – RR	Boa Vista	9,5	9,5	78,6	11,9	12,0	79,2
15 – PA	Belém	7,8	7,8	39,8	11,8	11,8	27,1
16 – AP	Macapá	10,2	10,3	74,9	13,3	13,5	36,2
17 – TO	Palmas	10,2	10,2	47,1	17,8	18,0	42,4
21 – MA	São Luís	5,4	5,5	33,5	9,2	9,3	23,0
22 – PI	Teresina	8,9	9,0	54,2	9,1	9,2	41,4
23 – CE	Fortaleza	4,6	4,6	..	6,0	6,0	..
24 – RN	Natal	7,4	7,4	..	9,1	9,1	..
25 – PB	João Pessoa	8,2	8,2	..	7,7	7,7	..
26 – PE	Recife	6,3	6,3	..	6,3	6,3	..
27 – AL	Maceió	4,5	4,5	33,5	12,0	12,0	19,7

Nota: Pela inexistência de setores rurais, em algumas capitais não há estimativas para este domínio.

Tabela 7 - Coeficiente de variação (%) estimado do estimador do total de desocupados e do estimador total da renda dos responsáveis por situação de setores segundo as capitais por ciclo de acumulação trimestral (conclusão)

UFs	Capitais	Desocupados			Renda		
		Total	Urbano	Rural	Total	Urbano	Rural
21 – MA	São Luís	5,4	5,5	33,5	9,2	9,3	23,0
22 – PI	Teresina	8,9	9,0	54,2	9,1	9,2	41,4
23 – CE	Fortaleza	4,6	4,6	..	6,0	6,0	..
24 – RN	Natal	7,4	7,4	..	9,1	9,1	..
25 – PB	João Pessoa	8,2	8,2	..	7,7	7,7	..
26 – PE	Recife	6,3	6,3	..	6,3	6,3	..
27 – AL	Maceió	4,5	4,5	33,5	12,0	12,0	19,7
28 – SE	Aracaju	8,0	8,0	..	9,7	9,7	..
29 – BA	Salvador	5,2	5,2	31,4	4,8	4,8	60,2
31 – MG	Belo Horizonte	6,4	6,4	..	8,6	8,6	..
32 – ES	Vitória	10,2	10,2	..	14,2	14,2	..
33 – RJ	Rio de Janeiro	4,7	4,7	..	4,2	4,2	..
35 – SP	São Paulo	5,1	5,5	6,6	5,4	5,5	6,7
41 – PR	Curitiba	6,8	6,8	..	5,4	5,4	..
42 – SC	Florianópolis	11,4	11,6	41,5	14,2	14,3	96,6
43 – RS	Porto Alegre	7,1	7,3	33,1	5,5	5,5	23,5
50 – MS	Campo Grande	7,7	7,7	76,8	9,6	9,6	79,3
51 – MT	Cuiabá	9,0	9,1	61,5	9,2	9,2	41,3
52 – GO	Goiânia	9,8	9,8	54,3	8,4	8,5	32,0

Nota: Pela inexistência de setores rurais, em algumas capitais não há estimativas para este domínio.

## Avaliação da adequação do tamanho da amostra à POF

A utilização planejada para a Amostra Mestra no Sistema Integrado de Pesquisas Domiciliares do IBGE visa à seleção de subamostras de setores ou de subamostras de domicílios com controle de sobreposição para diversas pesquisas. A estruturação do sistema que compete à utilização da Amostra Mestra pode ser esquematizada da seguinte forma:

### 1) Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC):

- todos os setores da Amostra Mestra
- amostra de domicílios

### 2) Pesquisa de Orçamentos Familiares (POF):

- subamostra de setores da Amostra Mestra

- amostra de domicílios não coincidente com a amostra da PNADC

### 3) Pesquisa de Orçamentos Familiares Simplificada (POFs):

- subamostra de setores
- amostra de domicílios coincidente com a amostra da PNADC (após última visita) ou
- amostra de domicílios não coincidente com as amostras das demais pesquisas

### 4) Economia Informal Urbana (ECINF):

- subamostra de setores
- amostra de domicílios coincidente ou não com a amostra das demais pesquisas

### 5) Pesquisas Especiais e Suplementares:

- subamostra de setores
- amostra de domicílios coincidente com a amostra da PNADC ( em alguma visita) ou
- amostra de domicílios coincidente com a amostra da POFs ou
- amostra de domicílios não coincidente com a amostra das demais pesquisas

Os dois primeiros pontos já estão definidos, os outros ainda estão em discussão. A Pesquisa de Orçamentos Familiares (POF 2008/2009) será a primeira pesquisa que fará uso da Amostra Mestra e tem como objetivo a divulgação de estimativas para a área urbana e rural separadamente, e o principal parâmetro de interesse é a renda, com isso foi preciso fazer uma avaliação dos tamanhos de amostra apresentados na tabela 4, para verificar se são adequados aos objetivos da POF.

Para verificar a adequação dos tamanhos de amostra da Amostra Mestra aos objetivos da POF 2008/2009, foi calculado o tamanho de amostra de setores necessário para estimar a renda total dos responsáveis pelos domicílios com a precisão desejada para a pesquisa e comparado com o tamanho da Amostra Mestra.

Observou-se que os tamanhos são compatíveis, ou seja, o tamanho da amostra de setores para a POF 2008/2009 é menor que os da Amostra Mestra, o que permitirá selecionar uma subamostra de setores. Apenas em alguns estratos foi preciso fazer um pequeno ajuste nos tamanhos para que fossem compatíveis.

Os tamanhos obtidos para a POF 2008/2009 são mostrados na tabela 8.

Tabela 8 - Tamanho da amostra de setores para estimar o total da renda dos responsáveis pelos domicílios e coeficiente de variação efetivo por situação do setor, segundo a UF - Pesquisa de Orçamentos Familiares 2008/2009

UF	Situação do setor					
	Total		Urbano		Rural	
	CV(%)	Número de setores	CV(%)	Número de setores	CV(%)	Número de setores
<b>Brasil</b>	<b>2,2</b>	<b>4 696</b>	<b>2,4</b>	<b>3 897</b>	<b>3,7</b>	<b>799</b>
<b>NORTE</b>	<b>5,8</b>	<b>601</b>	<b>6,5</b>	<b>459</b>	<b>11,7</b>	<b>142</b>
11 – RO	9,9	73	12,1	54	15,6	19
12 – AC	11,1	66	12,5	52	22,7	14
13 – AM	13,6	105	14,6	81	19,0	24
14 – RR	13,9	55	14,7	45	37,6	10
15 – PA	10,4	156	11,9	112	19,5	44
16 – AP	14,1	44	14,6	34	22,7	10
17 – TO	14,4	102	14,9	81	47,4	21
<b>NORDESTE</b>	<b>3,2</b>	<b>1 571</b>	<b>3,5</b>	<b>1 312</b>	<b>5,8</b>	<b>259</b>
21 – MA	8,9	209	10,6	172	13,3	37
22 – PI	10,2	153	12,0	123	8,9	30
23 – CE	10,1	143	11,2	103	9,3	40
24 – RN	9,4	113	9,7	95	33,5	18
25 – PB	8,6	128	9,6	107	11,4	21
26 – PE	7,2	193	7,7	162	15,1	31
27 – AL	11,1	246	12,3	226	21,7	20
28 – SE	9,3	141	9,9	135	25,7	6
29 – BA	6,8	245	7,5	189	13,8	56
<b>SUDESTE</b>	<b>3,6</b>	<b>1 234</b>	<b>3,8</b>	<b>1 071</b>	<b>7,8</b>	<b>163</b>
31 – MG	5,5	439	5,9	384	12,7	55
32 – ES	5,8	330	6,0	311	22,0	19
33 – RJ	7,2	171	7,3	145	21,2	26
35 – SP	5,4	294	5,6	231	12,2	63
<b>SUL</b>	<b>3,7</b>	<b>602</b>	<b>4,0</b>	<b>503</b>	<b>6,7</b>	<b>99</b>
41 – PR	6,3	231	6,8	198	12,4	33
42 – SC	6,2	182	6,8	152	13,8	30
43 – RS	6,0	189	6,5	153	9,4	36
<b>CENTRO-OESTE</b>	<b>4,7</b>	<b>688</b>	<b>5,0</b>	<b>552</b>	<b>13,5</b>	<b>136</b>
50 – MS	9,1	166	9,8	135	23,7	31
51 – MT	9,4	208	9,8	162	30,7	46
52 – GO	8,9	197	9,6	155	19,1	42
53 – DF	8,9	117	9,0	100	39,8	17

Nota: Os cálculos foram efetuados utilizando a base de dados do Censo Demográfico 2000.

## Seleção da Amostra

O método de seleção escolhido foi Amostragem de Pareto PPT (ver Costa, 2007) combinado com a técnica de números aleatórios permanentes, que associa a cada setor um valor,  $Q_i$ , obtido por uma expressão que considera um número aleatório entre 0 e 1, gerado para cada setor e que permanecerá atrelado a ele permanentemente, e o tamanho relativo do setor no grupo de rotação (número de domicílios do setor dividido pelo número de domicílios no grupo de rotação)<sup>2</sup>.

Após calcular para cada setor o  $Q_i$ , eles são ordenados por este valor em ordem crescente e são selecionados os  $m_g$  primeiros setores, onde  $m_g$  é o tamanho da amostra no grupo  $g$ .

Para a implementação do método de seleção foi preciso definir os grupos de rotação da amostra de domicílios da PNADC, pois para o controle da rotação da amostra de setores prevista (veja próxima seção), a seleção foi feita em cada grupo. Para isso os setores da população de cada estrato foram separados nos grupos de rotação antes da seleção. O procedimento de formação dos grupos de rotação e o procedimento de seleção são descritos a seguir.

### Definição dos grupos de rotação

Passos para definição dos grupos de rotação

- 1- Juntar os cadastros de setores de todos os estados em um arquivo Brasil
- 2- Excluir setores de tipos diferentes de Normal (1), Aglomerado subnormal (2) e Aldeias (5)
- 3- Contar o número de setores por estrato
- 4- Determinar o número de grupos que serão criados naquele estrato de acordo com o procedimento abaixo:
  - 4.1- Para os 16 estratos com 30 setores ou menos, será criado 1 grupo
  - 4.2- Para os 25 estratos que têm entre 31 e 105 setores, inclusive, serão criados 3 grupos
  - 4.3- Para os 509 estratos que têm mais que 106 setores serão criados 15 grupos
- 5- Criar os grupos nos estratos de duas maneiras diferentes:
  - 5.1- Nos estratos de renda criar os grupos de setores de maneira que eles sejam heterogêneos dentro deles e semelhantes entre si, da seguinte maneira:

---

<sup>2</sup> O valor de  $Q_i$  permanece inalterado no período em que o cadastro de seleção não sofre alterações. Quando o cadastro é atualizado, o tamanho relativo do setor é modificado, permanecendo inalterado o número aleatório entre 0 e 1.

**5.1.1-** Ordenar os setores dentro dos estratos pelo Código do Setor, desta maneira setores geograficamente próximos ficam perto no cadastro

**5.1.2-** Alocar os setores nos grupos de tal modo que o 1º setor fica no grupo 1, o 2º no grupo 2, ..., o 14º no grupo 14, o 15º no grupo 15, o 16º no grupo 1, o 17º no grupo 2, o 18º no grupo 3, assim por diante

**5.2-** Nos estratos somente geográficos e espaciais, ou seja sem estratificação de renda, criar os grupos utilizando algoritmos de otimização, buscando a formação de grupos parecidos entre si em relação à variável renda

**6-** Fazer uma aleatorização dos códigos dos grupos, para evitar que os primeiros grupos fiquem maiores que os últimos grupos:

**6.1-** Nos estratos com 1 e 3 grupos:

**6.1.1-** Separar em arquivo próprio

**6.1.2-** Classificar os setores por ordem crescente de código do estrato e aleatoriamente por código do grupo, mantendo os setores de mesmo grupo juntos

**6.1.3-** Atribuir ao 1º grupo o código 11, ao 2º grupo  $2 \times 11 \bmod 15 = 7$ , ao 3º grupo  $3 \times 11 \bmod 15 = 3$ , ao 4º grupo... Foi escolhido o número 11 por ser primo, desta maneira percorreu-se todos os 15 grupos e grupos do mesmo estrato ficaram em meses diferentes da pesquisa. Garantiu-se, também, que os grupos ficassem alocados relativamente distantes um do outro

**6.2-** Nos estratos com 15 grupos:

**6.2.1-** Separar em arquivo próprio

**6.2.2-** Classificar os setores por ordem crescente de código do estrato e aleatoriamente por código do grupo, mantendo os setores de mesmo grupo juntos

**6.2.3-** Atribuir ao primeiro grupo de cada estrato um código entre 1 e 15 aleatoriamente escolhido e numerar os grupos seguintes em ordem crescente aplicando o módulo 15

## **Procedimentos operacionais de seleção**

Passos para seleção da amostra

**1-** Adicionar a informação do tamanho de amostra por estrato ao cadastro

**2-** Contar o número de setores por grupo

**3-** Dividir o tamanho da amostra de cada estrato pelo número de grupos daquele estrato, ou seja, calcular o tamanho da amostra em cada grupo. A princípio a amostra de todos os

grupos é igual, mas no caso da divisão não ser exata, os grupos com mais setores serão também os com maior tamanho de amostra

4- Sortear um número aleatório entre 0 e 1 para cada setor. Este número será permanentemente atrelado a cada setor (*ALEAT*)

5- Definir a variável de tamanho, Número de Domicílios Particulares Permanentes Ocupados (*DPPO<sub>i</sub>*) de cada setor, limitando em 50 e 500 para evitar probabilidades muito baixas ou muito altas de seleção

6- Calcular o valor de  $p = \frac{DPPO_i}{\sum_{i \in grupo} DPPO_i}$ , a proporção de *DPPO<sub>i</sub>* de cada setor dentro dos grupos

7- Calcular o valor de *Q<sub>i</sub>* de cada setor de acordo com a fórmula:

$$Q_i = \frac{ALEAT - ALEAT \times p}{p - ALEAT \times p}$$

8- Ordenar os setores dentro de cada grupo de maneira crescente pelo *Q<sub>i</sub>* e selecionar para a Amostra Mestra os primeiros de cada grupo, de acordo com os tamanhos previamente calculados

9- Substituir um setor selecionado se este estiver esgotado, ou seja ter *DPPO<sub>i</sub>* menor que 20, pelo setor seguinte na ordem de seleção.

## Rotação da Amostra

Nessa seção são apresentadas algumas idéias sobre a atualização do cadastro de seleção da Amostra Mestra e dos procedimentos para realizar a rotação da amostra de setores.

### Rotação da amostra de setores

Uma proposta para o esquema de rotação da amostra de setores é fazer com que a amostra seja trocada inteiramente ao final de 10 anos. Para isso é necessário substituir 2,5% dos setores a cada trimestre. Como são 40 trimestres em 10 anos, ao final deste período 100% dos setores da amostra seriam trocados ( $40 * 2,5 = 100$ ).

Desta forma, a Amostra Mestra para um determinado ano será composta pelos setores selecionados para o primeiro trimestre mais os setores selecionados para serem os

substitutos nos 3 trimestres seguintes ( $3 * 2,5\% = 7,5\%$ ). Na versão atual, a Amostra Mestra será composta por 13760 setores, 12800 setores selecionados para o primeiro trimestre e 960 setores para dar conta do procedimento de rotação.

A rotação de setores será implementada aproveitando-se o procedimento de seleção adotado, que permite a troca controlada de unidades selecionadas. Na primeira rodada de seleção, após associar a cada setor o valor de  $Q_i$ , são selecionados os  $m_g$  primeiros setores, onde  $m_g$  é o tamanho da amostra no estrato  $g$ , como dito na seção 7.

Na segunda rodada de seleção, desloca-se o início da seleção para se obter o percentual de rotação desejado. Por exemplo, supondo que  $m_g = 10$  e que o percentual de rotação seja de 10%, a amostra da primeira rodada será composta pelos setores de ordem 1 até 10, na segunda rodada, quando pretende-se mudar 10%, ou seja, 1 setor, a amostra será composta pelos setores de ordem 2 até 11. Assim a coincidência das amostras será de 90% (setores de ordem 2 até 10) como desejado.

Este método permite a incorporação de atualizações no cadastro de seleção, evitando o congelamento da amostra no momento da seleção inicial, o que não permitiria acompanhar a evolução de crescimento dos setores.

Não havendo mudanças no cadastro de seleção ao longo do tempo, seria possível saber quais setores estariam entrando e saindo da amostra, em todos os trimestres, no momento da primeira seleção. Como é comum mudanças, especialmente nos tamanhos dos setores, acréscimo ou decréscimo no número de domicílios (medida de tamanho adotada na seleção), a incorporação destas mudanças a cada seleção proporcionaria melhorias nas estimativas, porém o conhecimento dos setores na amostra seria obtido com pouca antecedência.

### **Atualização do Cadastro Mestre**

Como dito, o método de seleção permite a incorporação de mudanças no cadastro e essa incorporação é desejável. Falta definir como e quando incorporar as atualizações.

A idéia é selecionar a amostra de setores a cada trimestre, promovendo a rotação desejada de 2,5% da amostra. As atualizações se forem incorporadas a cada seleção, ocasionará um problema de ordem prática. Só se saberia quais setores estariam entrando na amostra com uma antecedência muito pequena, o que levaria a dificuldades de planejamento da coleta. E se a incorporação das atualizações for feita com uma periodicidade grande, não trará grandes melhorias. Portanto o “ideal” seria um meio-termo.

A proposta pensada é levantar as atualizações continuamente, consolidá-las uma vez por ano, incorporando-as ao cadastro, e selecionar em julho a Amostra Mestra correspondente aos 4 trimestres do ano seguinte. Desta forma seriam conhecidos os novos setores a serem visitados com 6 meses de antecedência no mínimo. Como o cadastro será “congelado” uma vez por ano, ao selecionar a amostra do primeiro trimestre, já se pode

obter os setores que entrarão nos trimestres subsequentes, utilizando o método de seleção e rotação descritos.

Para a seleção da Amostra Mestra a informação atualizada importante é o **total de domicílios no setor**, que deverá ser incorporada ao Cadastro Mestre (que contém TODOS os setores censitários, os que estão na amostra e os que não estão). O **total de domicílios**, informação básica, necessária e suficiente e que deve ser atualizada, poderá ser obtida de diversas fontes: das pesquisas do SIPD, das listagens de domicílios, de fontes administrativas, do projeto do Cadastro Nacional de Endereços para fins estatísticos do IBGE e de outras fontes que forem pensadas e descobertas no futuro.

## Conclusões

O desenho da amostra da Amostra Mestra visa atender várias pesquisas, buscando satisfazer os seus diversos objetivos, e com isso é de se esperar que o desenho amostral não seja o melhor para cada pesquisa isoladamente, mas seja bom para todas.

No atual estágio de planejamento do Sistema Integrado de Pesquisas Domiciliares do IBGE, foram selecionadas a Amostra Mestra, a subamostra de setores para a Pesquisa de Orçamentos Familiares, prevista para ir a campo em 2008/2009 e a subamostra de setores da Pesquisa sobre Características Étnico-raciais da População, que deverá ter os dados coletados em março de 2008.

Ainda se torna necessário definir a forma do cadastramento de domicílios (listagem) que será utilizado para seleção da amostra nas diversas pesquisas, bem como o método de seleção. Cabe ressaltar que deverá ser criada uma identificação permanente do domicílio no cadastro, para coordenação das amostras, controlando a sobreposição desejada.

## Referências Bibliográficas

ANTONACI, G.; SILVA, D.B.N. Emparelhamento de domicílios e pessoas na Pesquisa Mensal de Emprego e cálculo da autocorrelação da característica desocupação. In: SIMPÓSIO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 17. 2006, Caxambu. *Resumos*. São Paulo: Associação Brasileira de Estatística, 2006.

AZEVEDO, R.V. *Estudo comparativo de métodos de estratificação ótima de populações assimétricas*. 2004. Dissertação (Mestrado) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro.

ASSUNÇÃO, R.M. *Análise de conglomerados espaciais para uso em amostragem estratificada com probabilidade de seleção proporcional ao tamanho*. Relatório final - BRA/97/013. Projeto: Rede de pesquisas e desenvolvimento de políticas econômicas - Sistema Rede IPEA. Projeto 6 - Desenvolvimento e absorção de novas tecnologias de produção de informações, 2000.

BIANCHINI, Z.M.; ALBIERI, S. E. *Principais aspectos da amostragem da pesquisas domiciliares do IBGE – revisão 2002*. Rio de Janeiro: IBGE, Departamento de Metodologia. 2003. (Texto para discussão nº 8).

BUSSAB, W.O.; DINI. N.P. Pesquisa de emprego e desemprego SEADE/DIEESE: regiões homogêneas da Grande São Paulo. *Revista Fundação SEADE/São Paulo em perspectiva*, set/dez, p.5-11, 1985.

CARRILHO, A.; NTHABISENG, M. Sampling, weighting and standard error estimation methodology for the labour force survey conducted in September 2001. SACD workshop on Survey Sample Designs. Windhoek, Namibia, July 15-19, 2002.

COCHRAN, W.G. *Sampling Techniques*. 3rd. ed. New York: John Wiley, 1977.

COSTA, G.T.L. *Coordenação de amostras PPT em Pesquisas Repetidas, utilizando o método de amostragem de Pareto*. 2007. Dissertação (Mestrado) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro.

FREITAS, M.P.S.; LILA M.F. *Uma proposta de dimensionamento de amostra para a pesquisa domiciliar contínua (versão preliminar)*. Rio de Janeiro: IBGE, Coordenação de Trabalho e Rendimento. 2004.

FREITAS, M.P.S. *Estratificação para amostra de uma pesquisa domiciliar sobre mercado de trabalho*. 2002. Dissertação (Mestrado) – Escola Nacional de Ciências Estatísticas, Rio de Janeiro.

KISH, L. *Survey Sampling*. Nova Iorque: John Wiley & Sons, 1965

LILA M.F.; FREITAS, M.P.S. *Estimação de intervalos de confiança para estimadores de diferenças temporais na Pesquisa Mensal de Emprego*. Rio de Janeiro: IBGE, 2007. (Texto para discussão, nº 22)

LIMA, M.I.F.; BIANCHINI, Z.M. Estudos para o dimensionamento da amostra da pesquisa de orçamentos familiares. In: SIMPÓSIO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 7. 1986, Campinas. *Resumo*. São Paulo: Associação Brasileira de Estatística, 1988.

MONTENEGRO, F.; BRITO, J. A. M. . Um Algoritmo Genético para o Problema de Agrupamento. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 38. 2006, Goiânia. Anais do XXXVIII SOBRAPO, 2006.

MLADENOVIC, N.; HANSEN, P. Variable neighborhood search: Principles and applications. *European Journal of Operational Research*, v. 10, p. 449-467, 1999.

PETERSSON, H. *Household Sample Surveys in Developing and Transition Countries: Design of master sampling frames and master samples for household surveys*. Chapter V, United Nations Statistics Division, 2005.

RESENDE, M.G.C.; SOUSA, J.P. *Metaheuristics: Computer Decision-Making*. Kluwer Academic Publishers. 2004.

SÄRNDAL, C.E.; SWENSSON, B.; WRETMAN, J.H. Model assisted survey sampling. New York: Spring-Verlag. (1992).

SILVA, P.L.N. *Algumas idéias para a revisão das pesquisas domiciliares por amostragem do IBGE*. [Rio de Janeiro]: IBGE, Departamento de Metodologia, 2001.

SILVA, P.L.N. et al. *Aspectos sobre a estrutura longitudinal no contexto da Pesquisa sobre Mercado de Trabalho*. [Rio de Janeiro]: IBGE, Departamento de Metodologia, 1998.

TURNER, A. G. Sampling frames and master sample, United Nations Statistics Division, 2003.

YANSANEH, S.; FULLER, W.A. Optimal Recursive Estimation for Repeated Surveys. *Survey Methodology*, Ottawa, volume 24, number 1, June 1998.

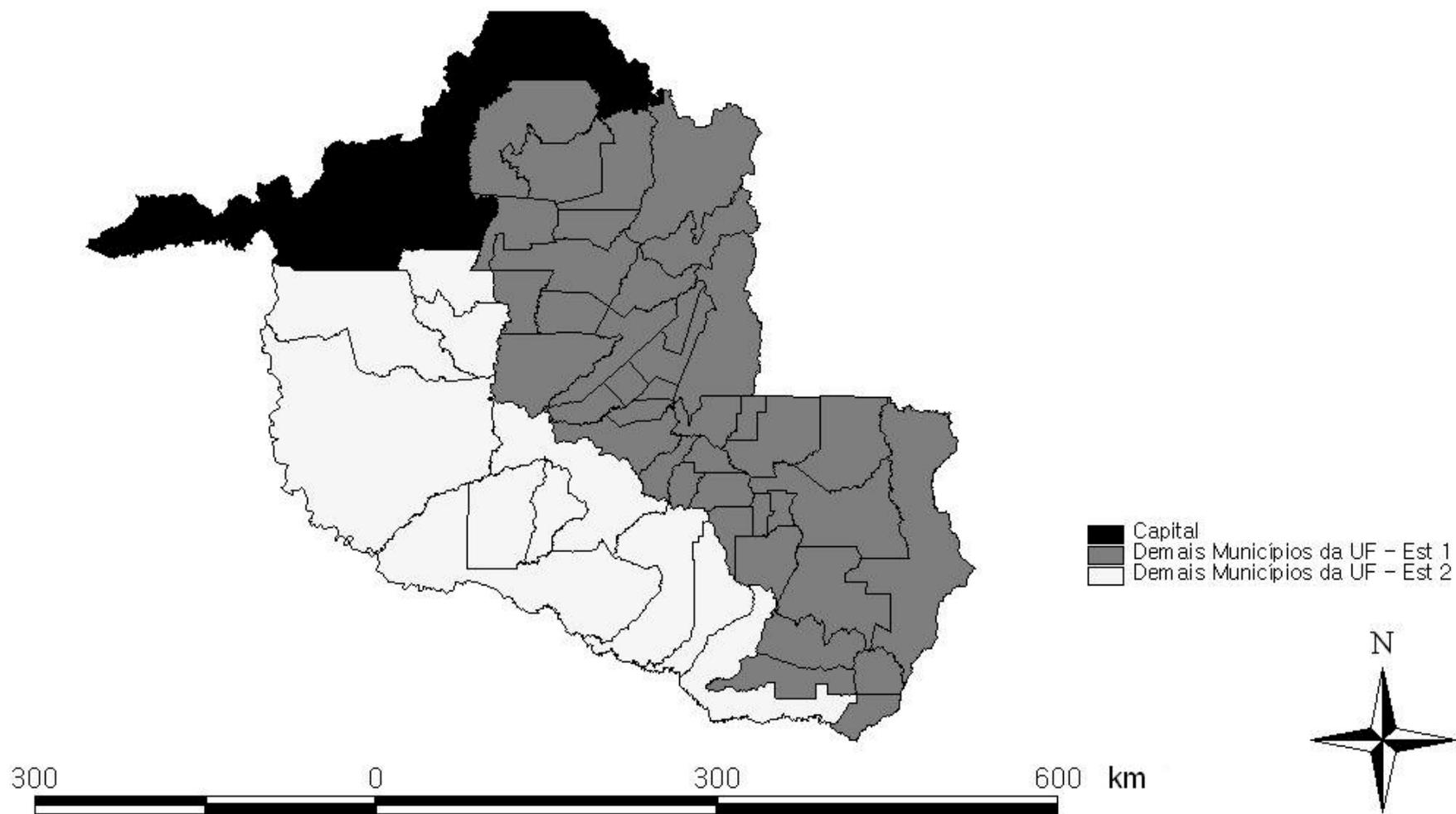
## Anexo 1 - Estratificação Geográfica: Tabelas e Mapas

Tabela A1 - Distribuição do número de setores, domicílios e pessoas segundo a UF

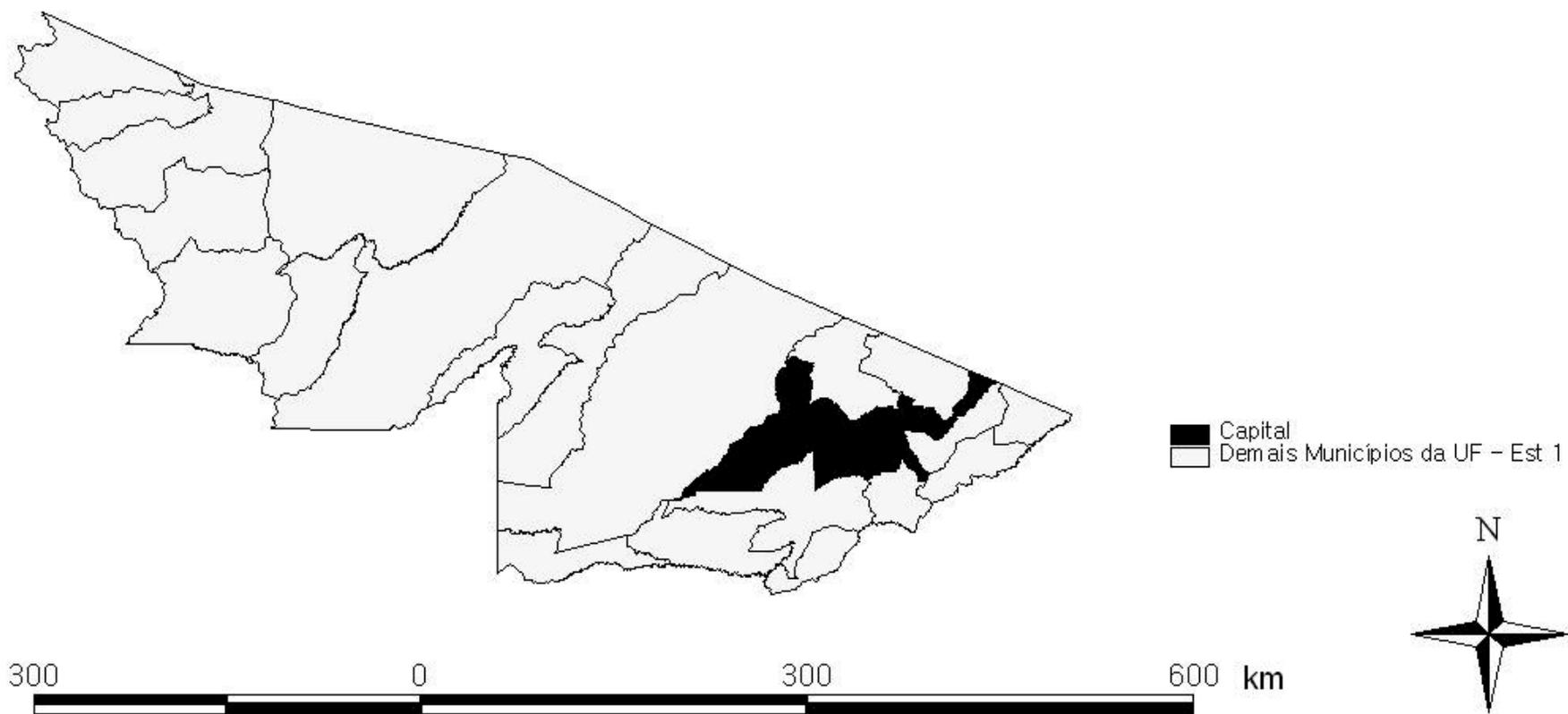
UF	Distribuição por Estrato Geográfico e Espacial											
	Setores			Total de Domicílios (DTOT)			Domicílios particulares permanentes ocupados abertos (DPPO)			Pessoas		
	Mínimo	Mediana	Máximo	Mínimo	Mediana	Máximo	Mínimo	Mediana	Máximo	Mínimo	Mediana	Máximo
<b>Brasil</b>	<b>67</b>	<b>963</b>	<b>13 899</b>	<b>18 091</b>	<b>241 521</b>	<b>3 618 818</b>	<b>15 794</b>	<b>203 427</b>	<b>2 773 925</b>	<b>79 766</b>	<b>830 068</b>	<b>9 920 501</b>
<b>11</b>	275	420	1 190	53 615	100 354	259 880	44 986	83 682	218 526	184 951	332 240	843 641
<b>12</b>	245	276	307	76 661	77 649	78 636	64 078	64 720	65 361	251 573	275 815	300 056
<b>13</b>	443	761	795	65 552	142 697	198 542	59 823	123 092	171 146	333 600	668 992	723 769
<b>14</b>	205	233	261	32 631	46 702	60 773	25 698	37 226	48 753	118 072	158 723	199 373
<b>15</b>	468	1 105	3 399	139 153	295 241	810 606	119 953	250 873	687 334	511 886	1 142 824	3 319 427
<b>16</b>	67	148	232	18 091	27 829	72 678	15 794	22 382	60 400	79 766	110 244	281 510
<b>17</b>	105	298	663	27 161	81 195	149 818	23 006	67 519	122 238	94 879	280 067	487 802
<b>21</b>	102	1 059	2 057	34 053	247 124	440 009	29 399	213 821	376 151	129 454	968 847	1 742 684
<b>22</b>	198	702	1 243	41 347	193 741	241 858	35 618	159 273	206 004	159 617	661 304	902 252
<b>23</b>	537	878	1 487	149 131	242 853	396 352	122 212	201 934	325 875	477 011	833 369	1 456 085
<b>24</b>	330	404	593	96 302	123 294	209 798	76 758	97 017	177 783	316 531	414 303	708 596
<b>25</b>	340	711	2 389	102 939	182 726	569 191	84 492	153 098	458 690	345 904	612 438	1 855 709
<b>26</b>	1 203	1 587	2 544	341 057	424 587	621 899	281 801	376 022	506 431	1 199 256	1 413 130	1 985 329
<b>27</b>	164	571	795	55 833	158 284	238 958	43 064	133 711	199 734	189 237	575 784	832 663
<b>28</b>	499	801	915	143 126	179 787	221 987	116 689	141 819	178 227	459 556	567 128	743 508
<b>29</b>	614	1 323	6 019	143 166	384 874	1 422 671	112 880	313 370	1 156 592	490 467	1 278 688	4 912 817
<b>31</b>	124	2 461	4 687	28 779	652 828	1 155 737	23 597	542 921	958 374	88 181	1 980 286	3 549 104
<b>32</b>	268	600	1 062	100 623	175 417	396 013	85 558	145 106	315 142	290 880	557 019	1 140 582
<b>33</b>	1 268	1 900	4 044	309 398	488 539	1 033 536	169 134	420 530	899 839	584 298	1 323 937	2 978 513
<b>35</b>	2 344	3 928	13 899	546 221	1 035 406	3 618 818	427 841	877 171	2 773 925	1 576 185	3 059 447	9 920 501
<b>41</b>	1 024	1 575	4 869	265 482	406 624	1 109 269	232 443	331 611	952 951	727 839	1 221 110	3 339 172
<b>42</b>	384	1 113	3 582	122 632	266 029	984 864	102 657	236 064	814 593	339 063	830 811	2 901 817
<b>43</b>	863	3 352	4 036	205 998	699 400	879 693	172 399	581 560	746 996	449 415	1 966 471	2 491 282
<b>50</b>	455	669	914	97 570	176 991	220 476	80 364	147 730	187 079	317 930	520 880	686 796
<b>51</b>	311	417	1 275	55 585	94 151	300 520	44 286	78 986	250 341	169 950	302 833	938 572
<b>52</b>	526	968	1 566	148 707	301 856	386 480	118 065	249 206	313 965	439 752	902 483	1 087 579
<b>53</b>	338	434	1 766	74 815	115 162	439 699	64 126	95 392	388 138	232 419	312 478	1 490 327

Fonte: Base Operacional Geográfica de 2000 compatibilizada com a malha municipal de 2001 e Censo 2000

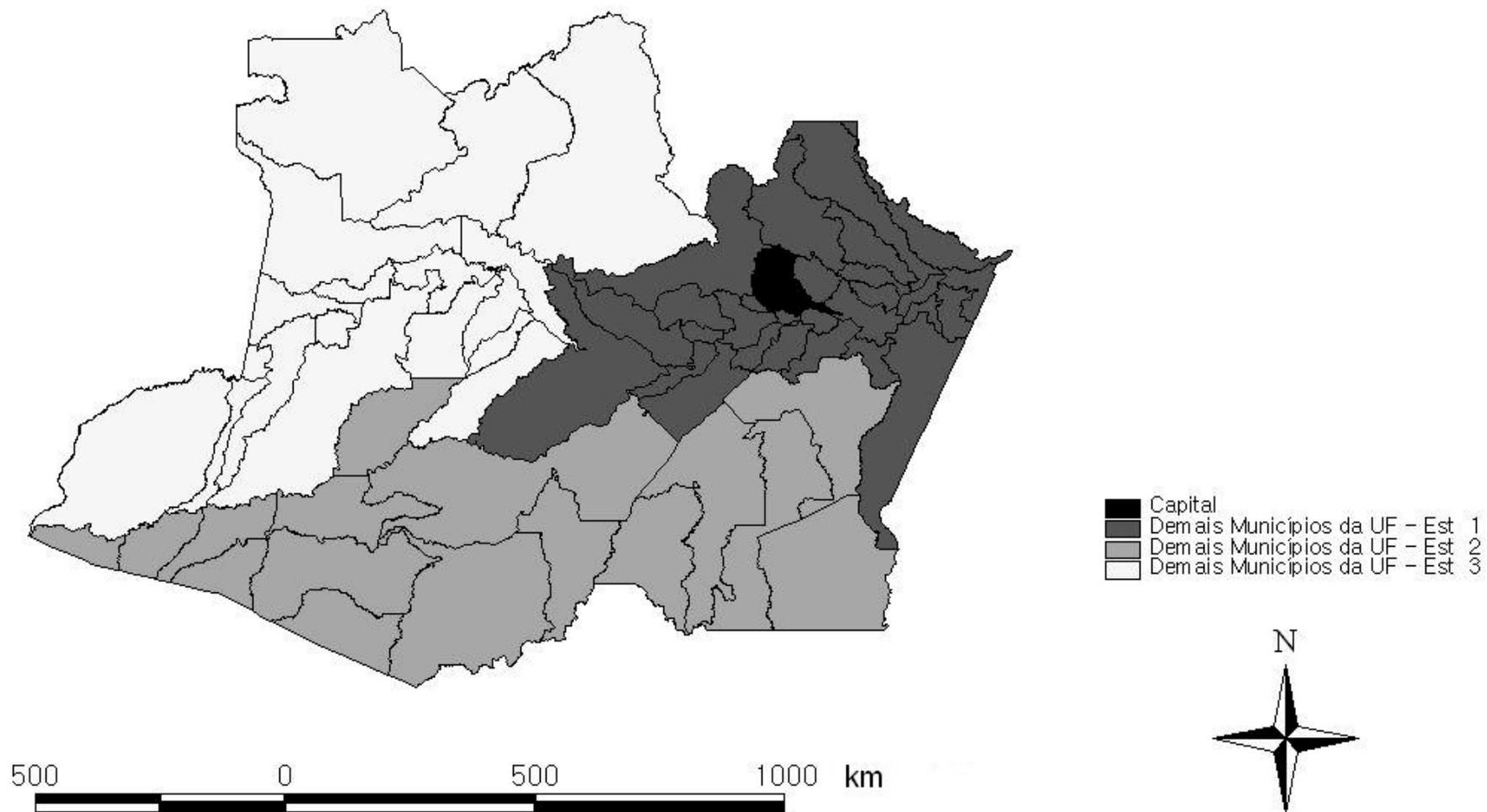
Mapa A1.1 - Distribuição dos municípios por estrato geográfico - Rondônia



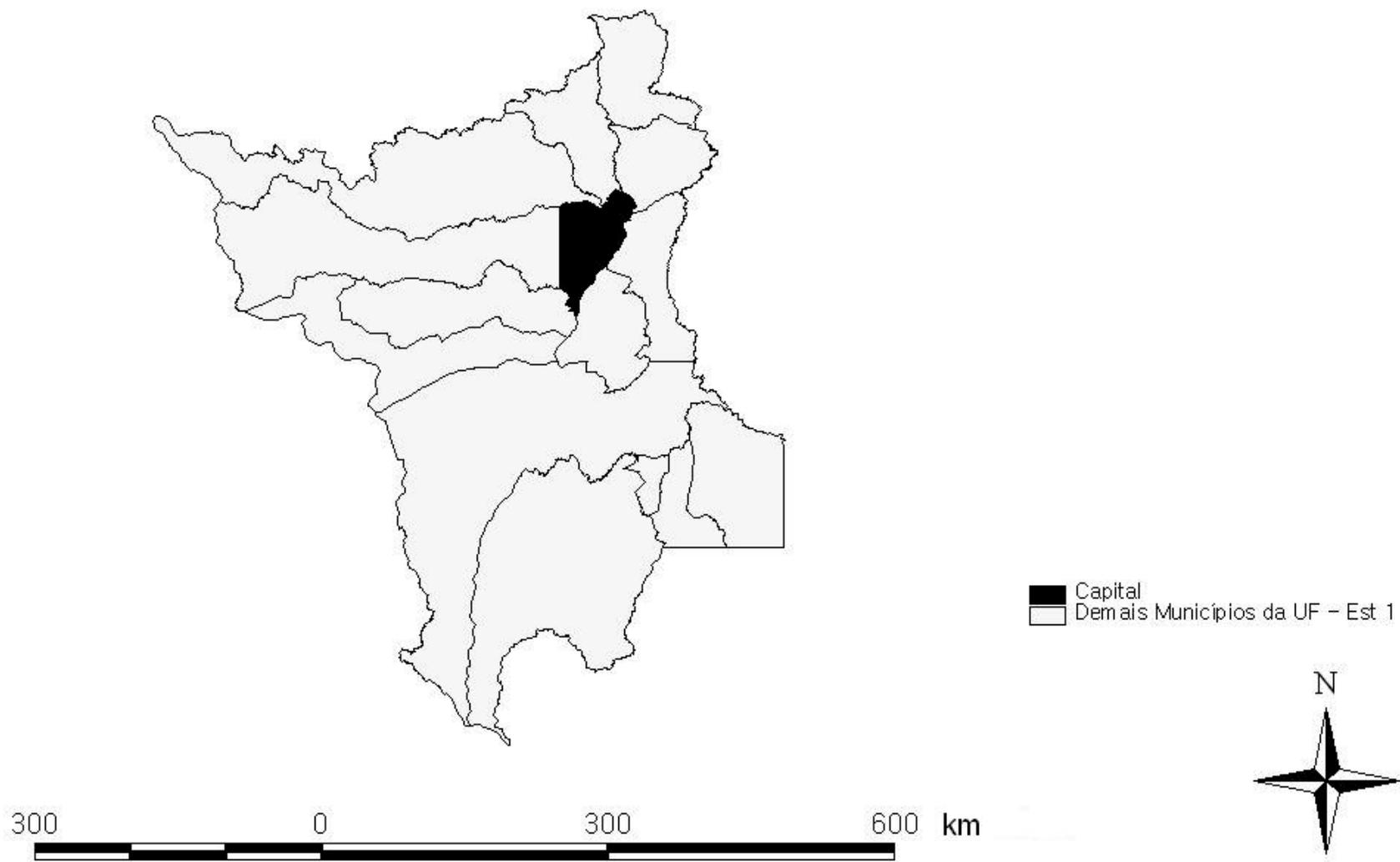
Mapa A1.2 - Distribuição dos municípios por estrato geográfico - Acre



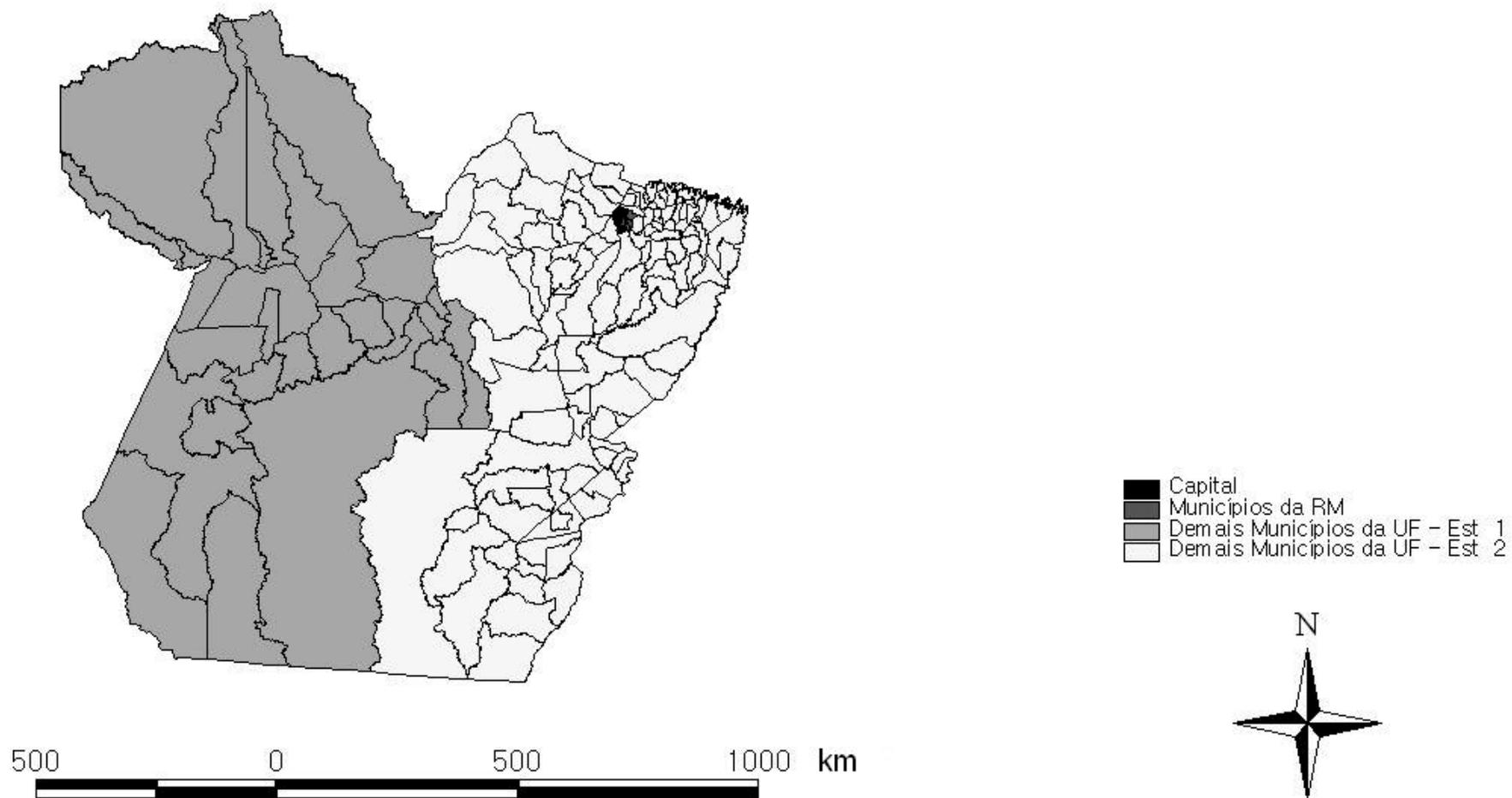
Mapa A1.3 - Distribuição dos municípios por estrato geográfico e espacial - Amazonas



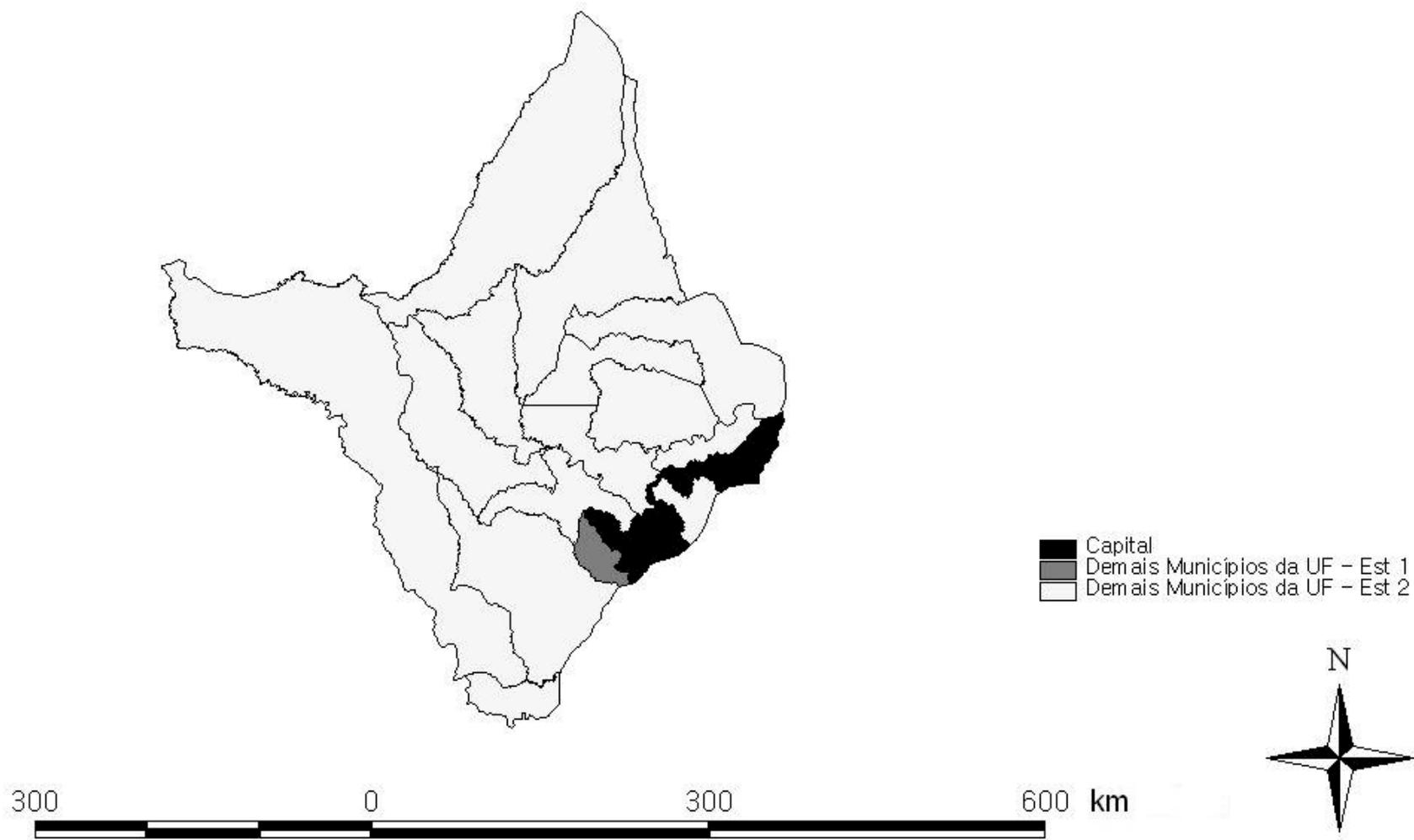
Mapa A1.4 - Distribuição dos municípios por estrato geográfico - Roraima



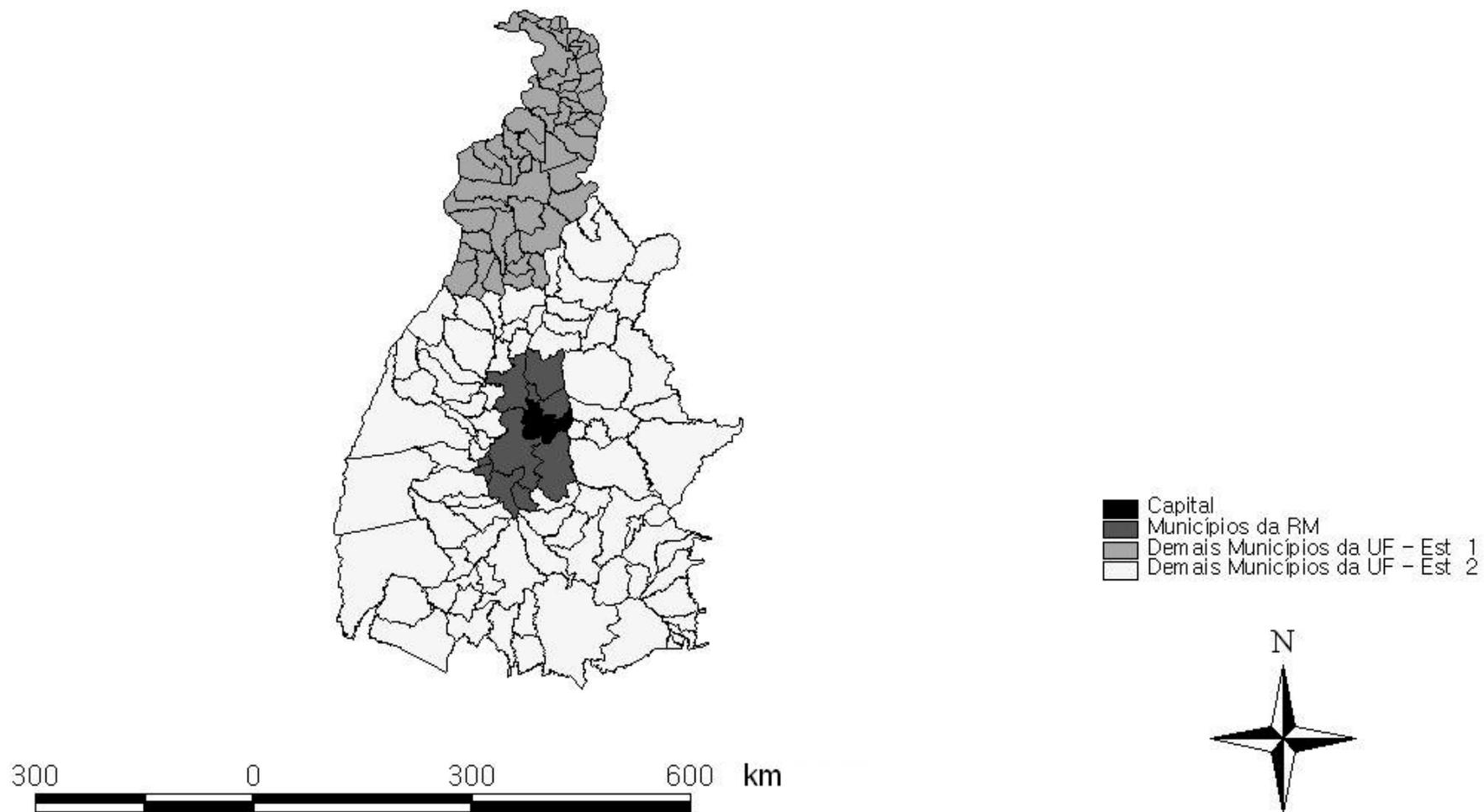
Mapa A1.5 - Distribuição dos municípios por estrato geográfico e espacial - Pará



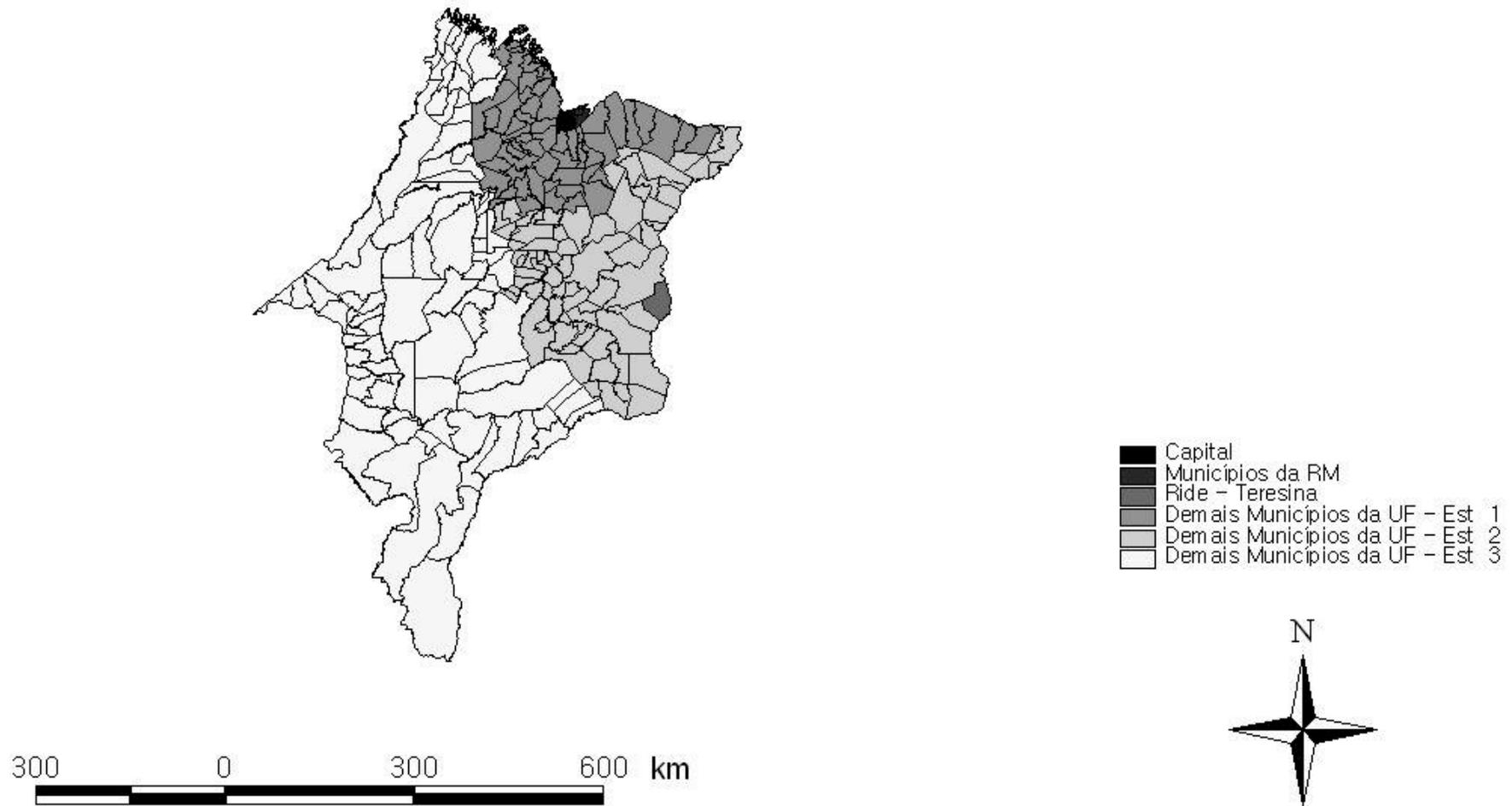
Mapa A1.6 - Distribuição dos municípios por estrato geográfico - Amapá



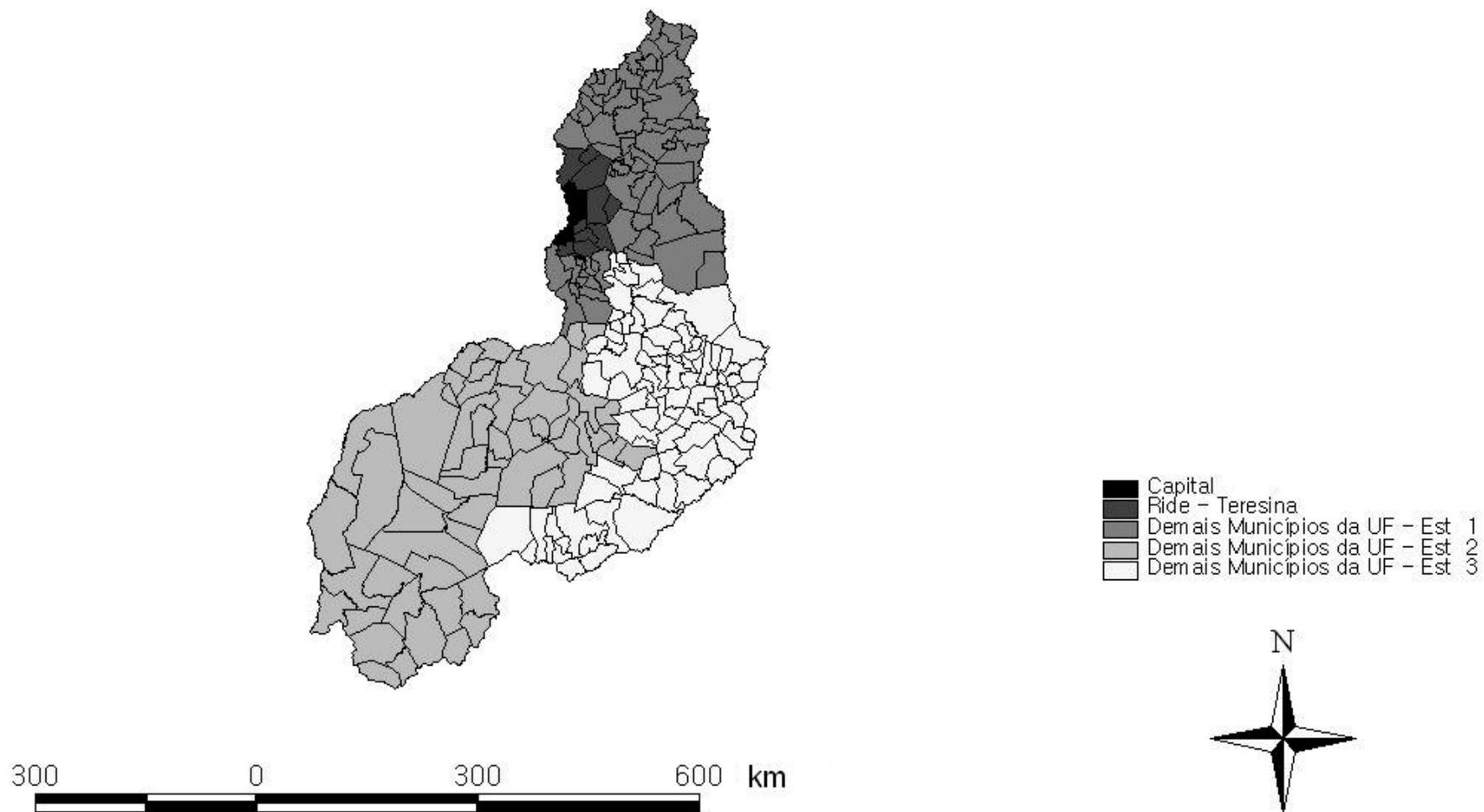
Mapa A1.7 - Distribuição dos municípios por estrato geográfico e espacial - Tocantins



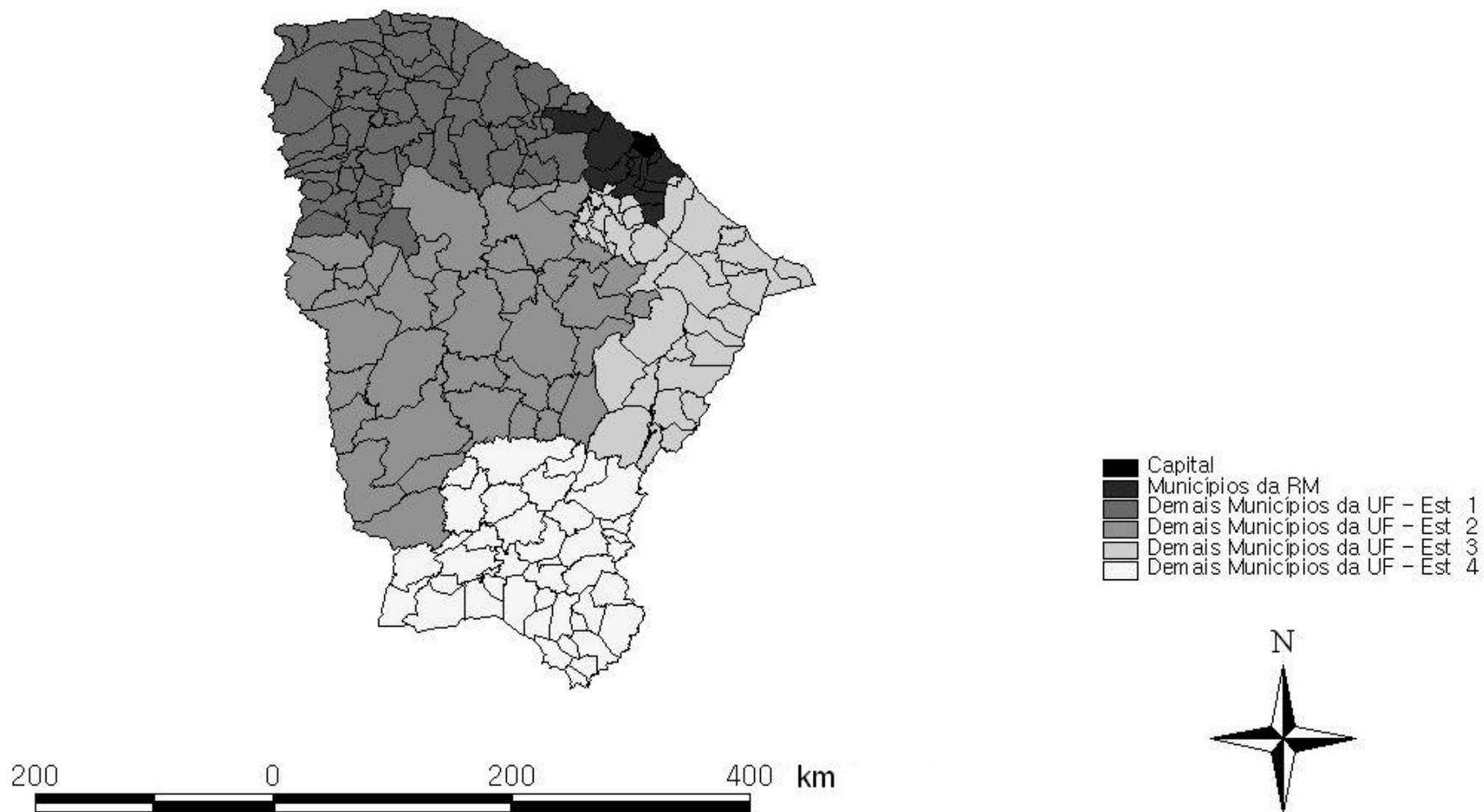
Mapa A1.8 - Distribuição dos municípios por estrato geográfico e espacial - Maranhão



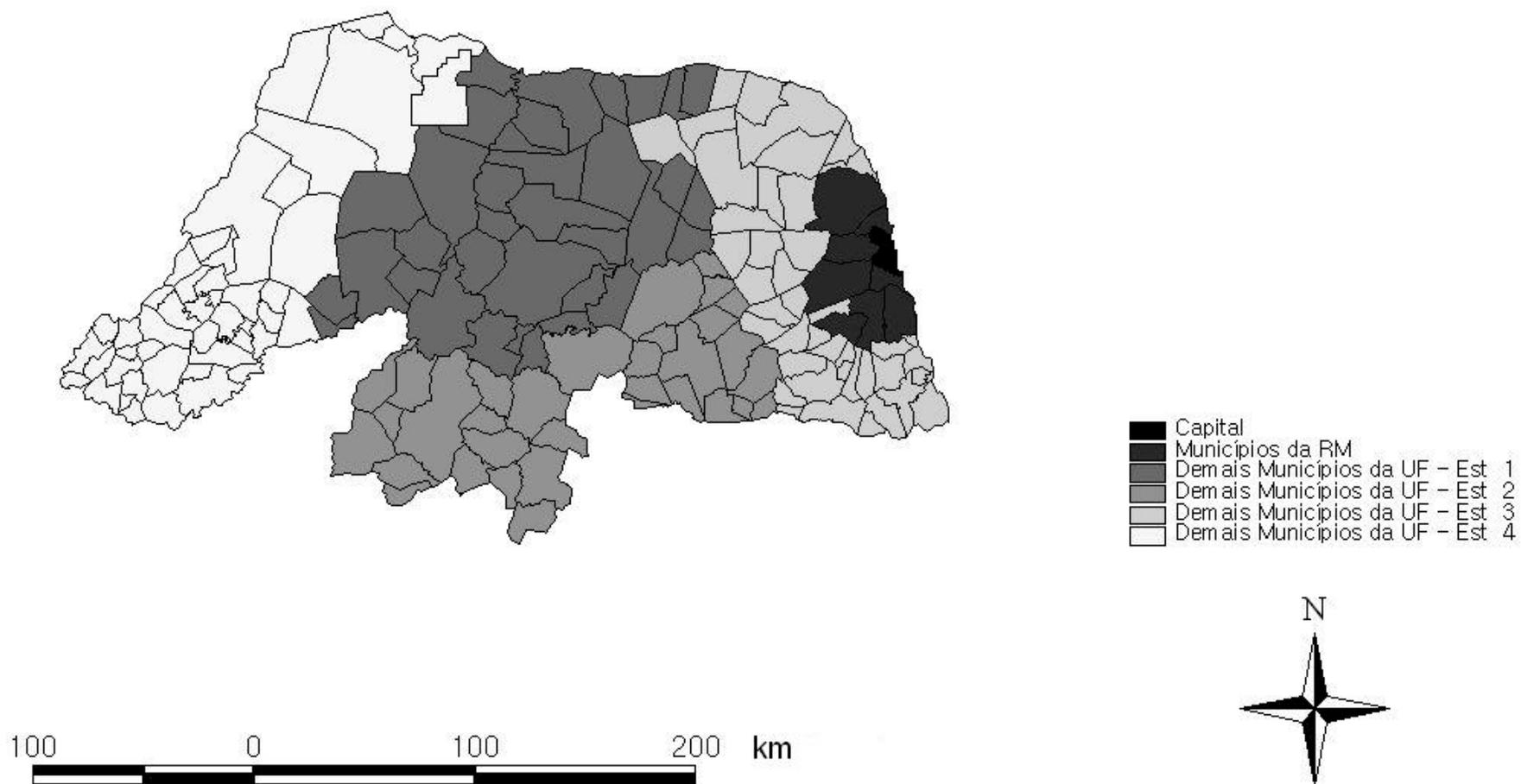
Mapa A1.9 - Distribuição dos municípios por estrato geográfico e espacial - Piauí



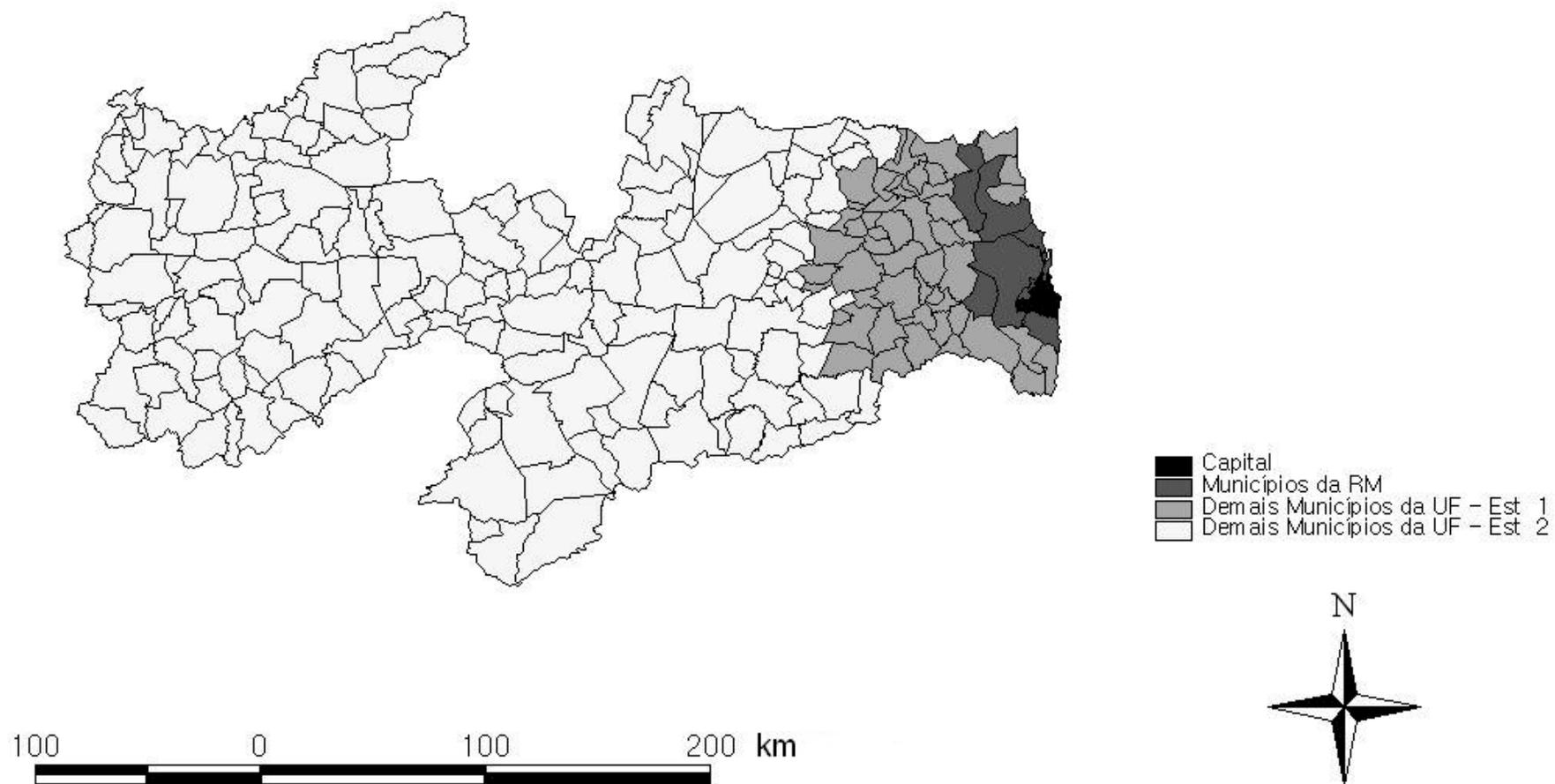
Mapa A1.10 - Distribuição dos municípios por estrato geográfico e espacial - Ceará



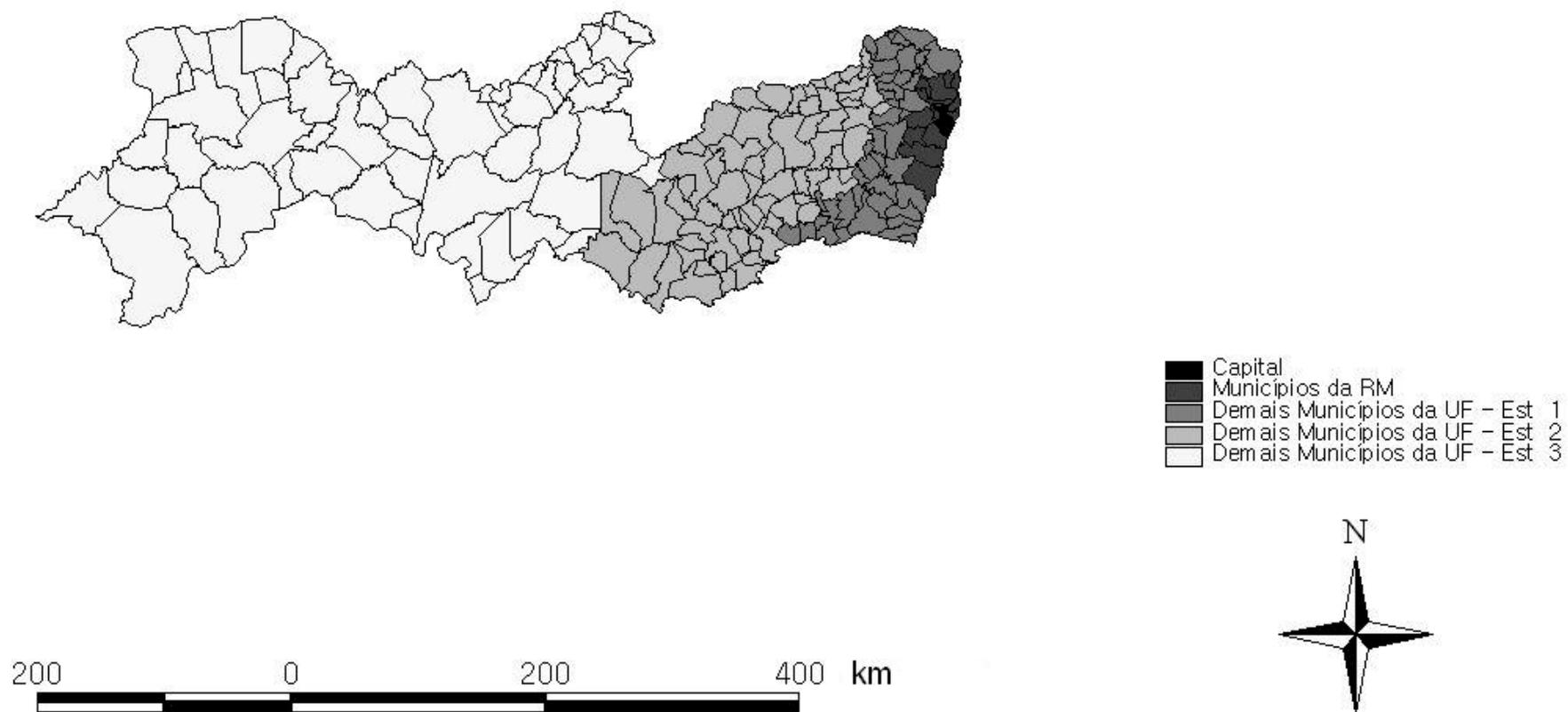
Mapa A1.11 - Distribuição dos municípios por estrato geográfico e espacial - Rio Grande do Norte



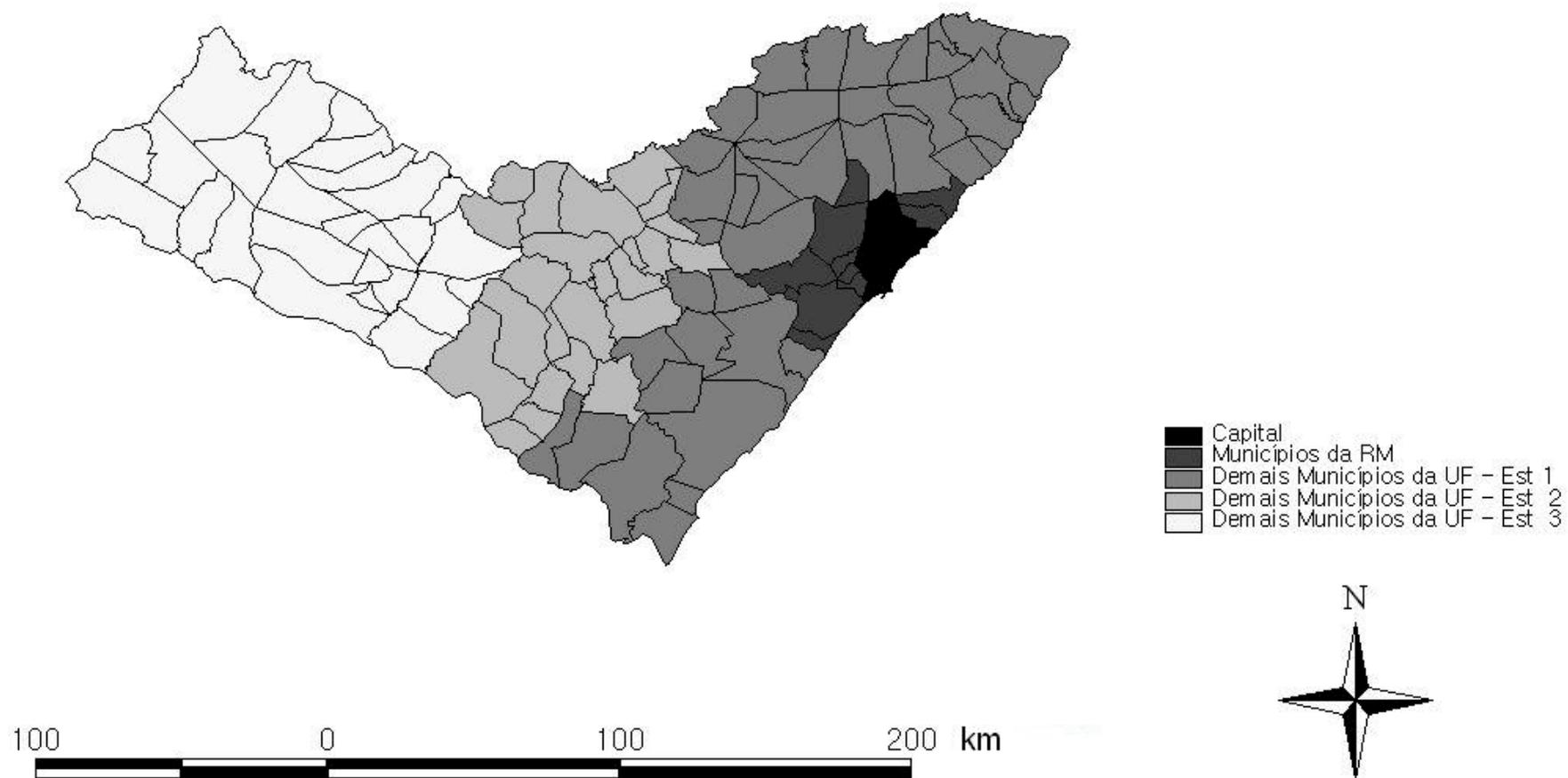
Mapa A1.12 - Distribuição dos municípios por estrato geográfico e espacial - Paraíba



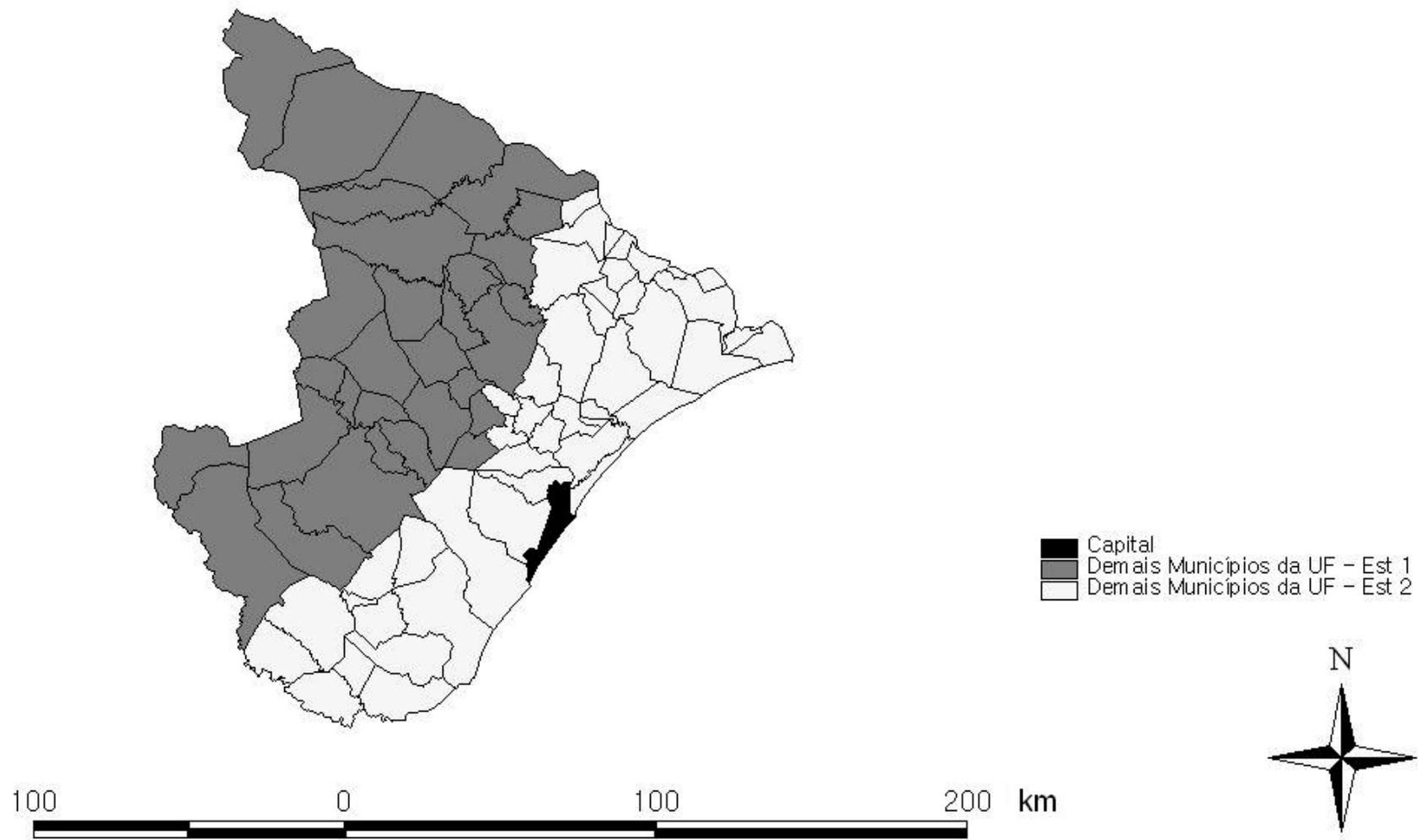
Mapa A1.13 - Distribuição dos municípios por estrato geográfico e espacial - Pernambuco



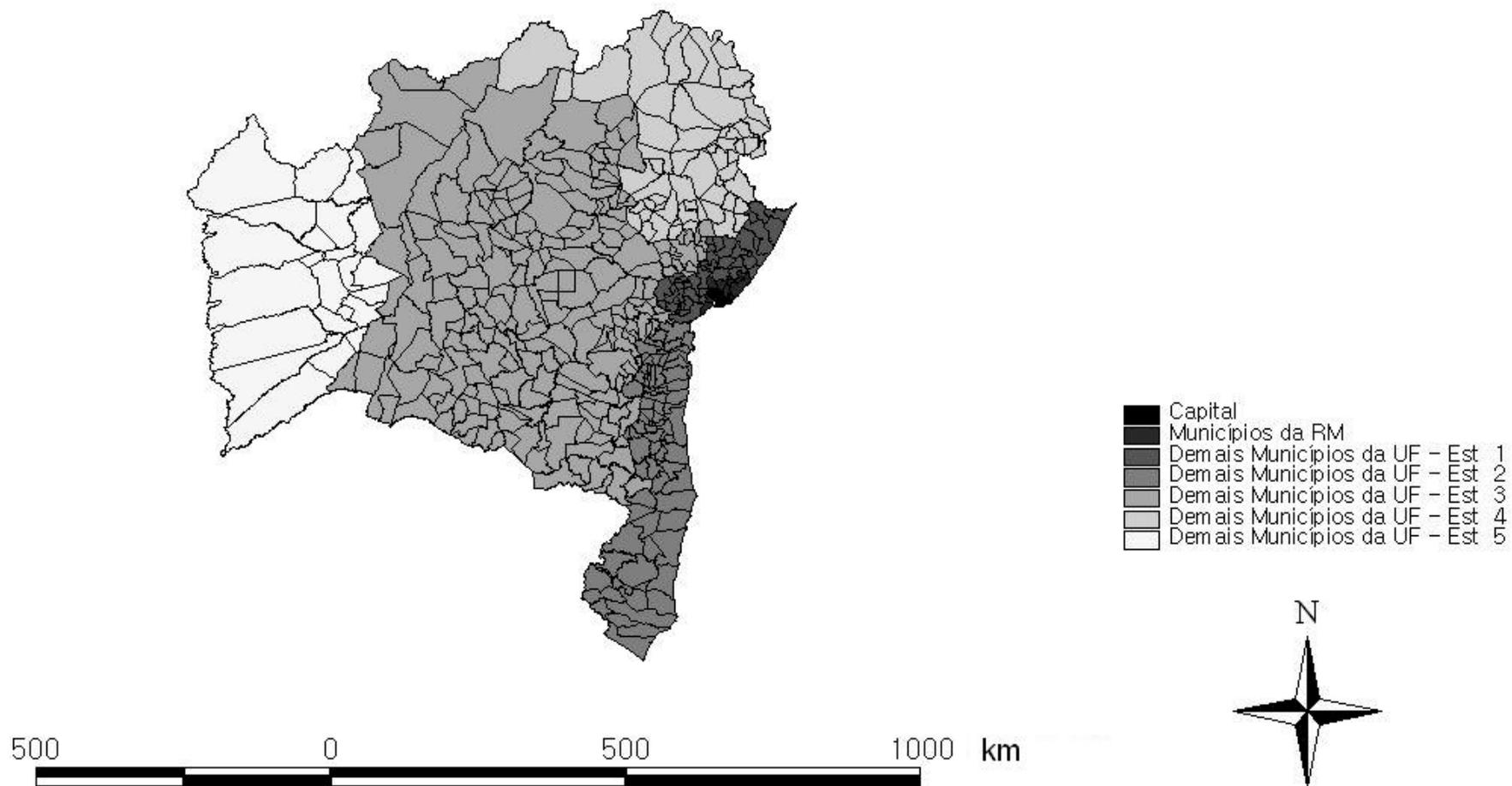
Mapa A1.14 - Distribuição dos municípios por estrato geográfico e espacial - Alagoas



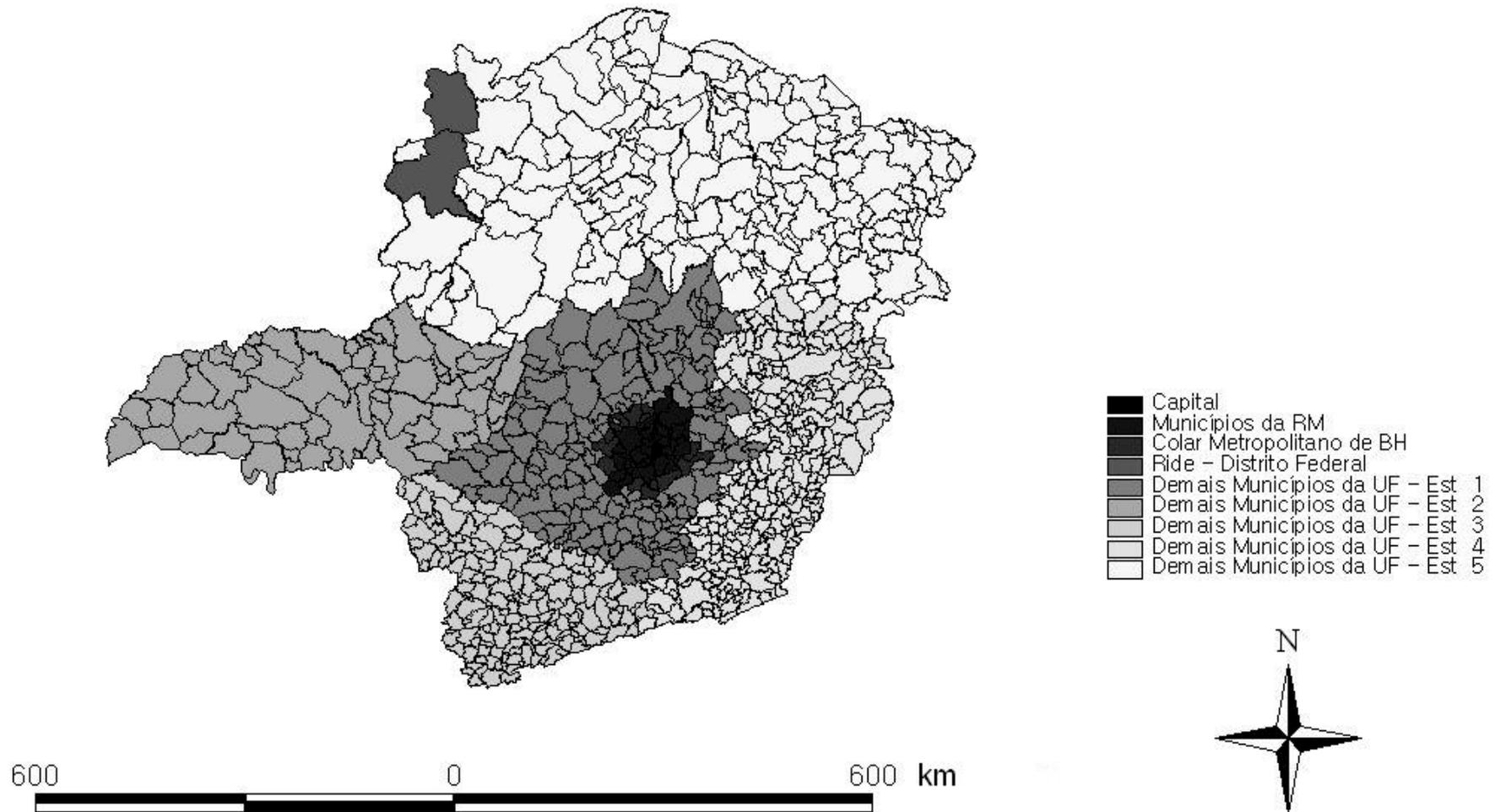
Mapa A1.15 - Distribuição dos municípios por estrato geográfico e espacial - Sergipe



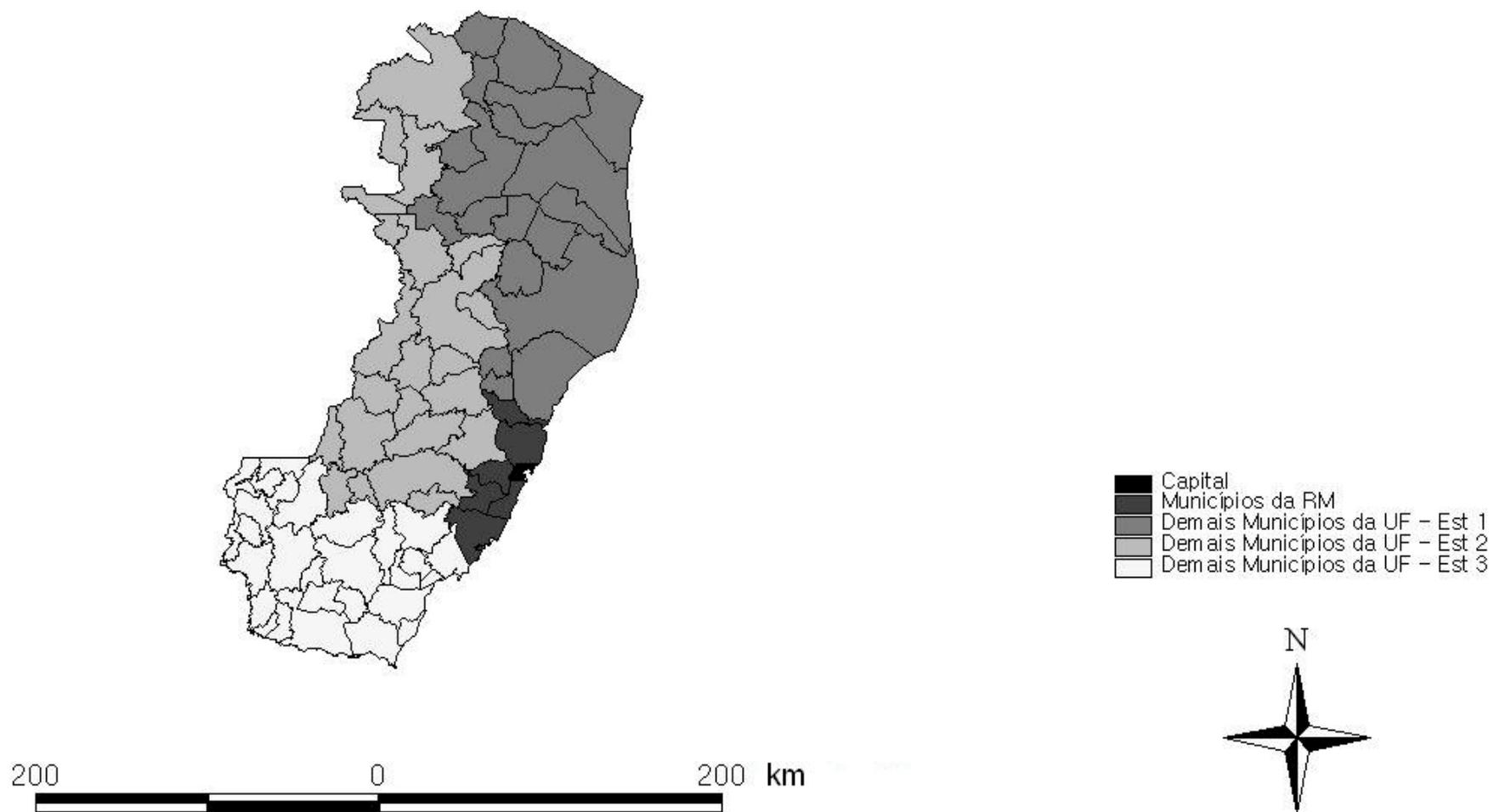
Mapa A1.16 - Distribuição dos municípios por estrato geográfico e espacial - Bahia



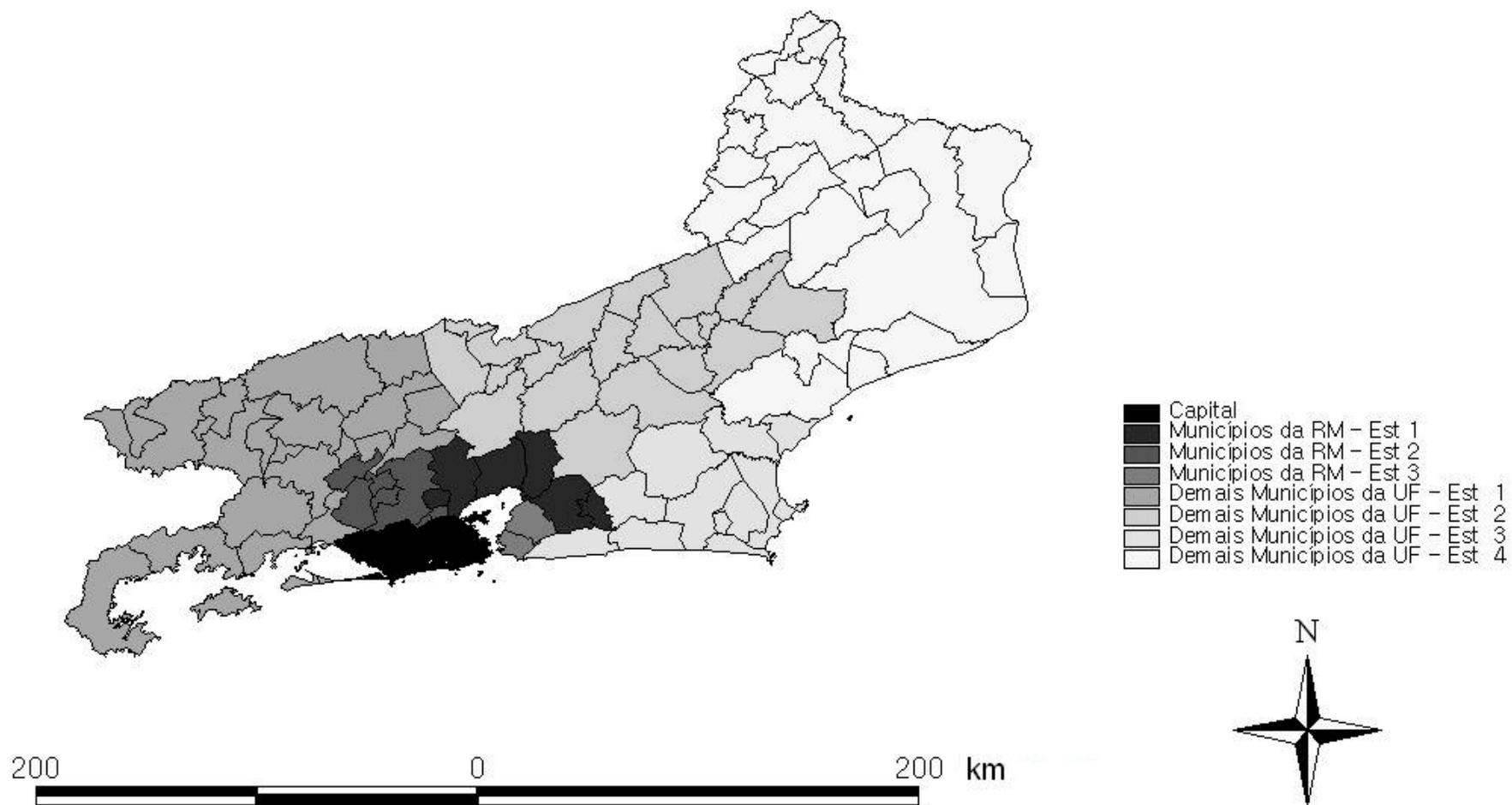
Mapa A1.17 - Distribuição dos municípios por estrato geográfico e espacial - Minas Gerais



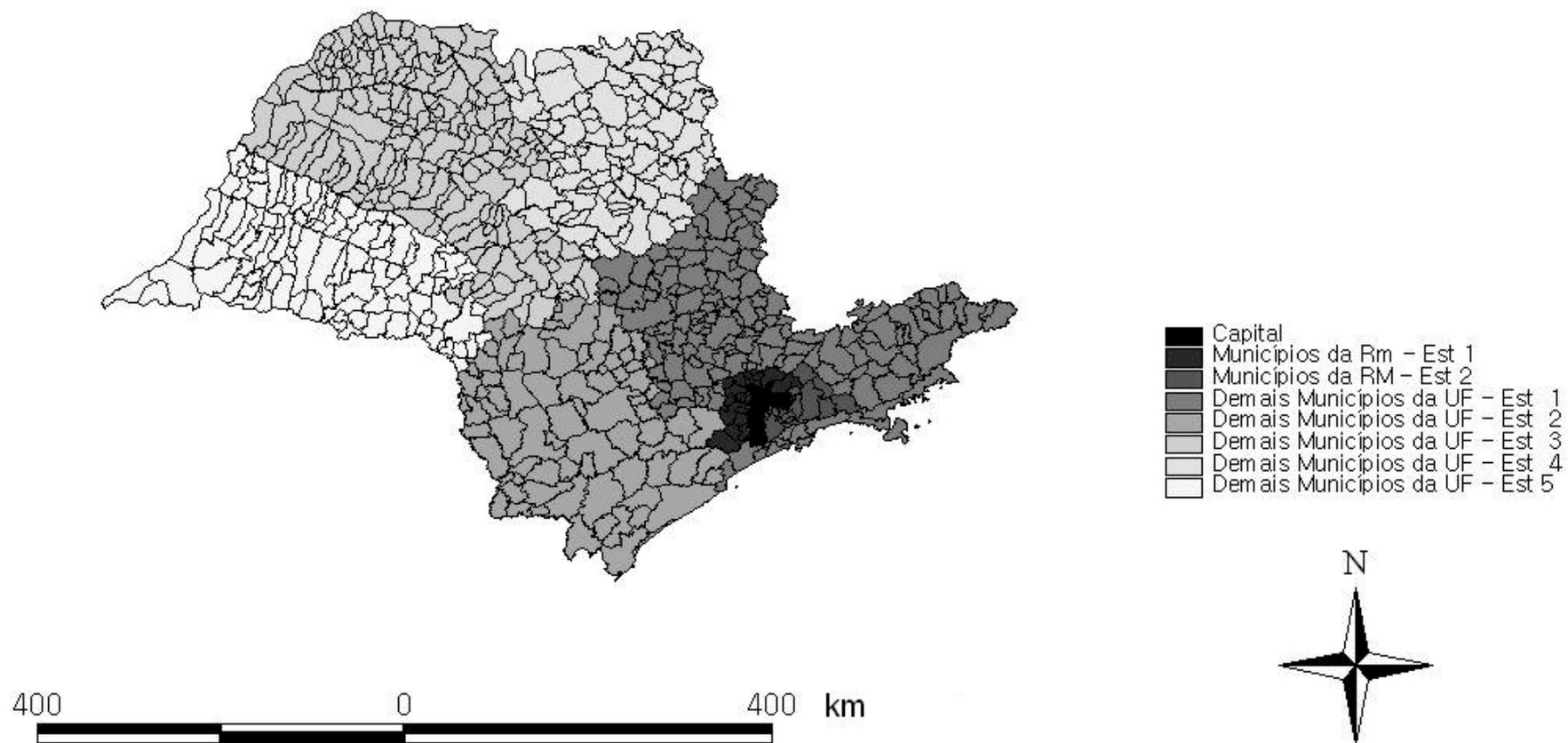
Mapa A1.18 - Distribuição dos municípios por estrato geográfico e espacial - Espírito Santo



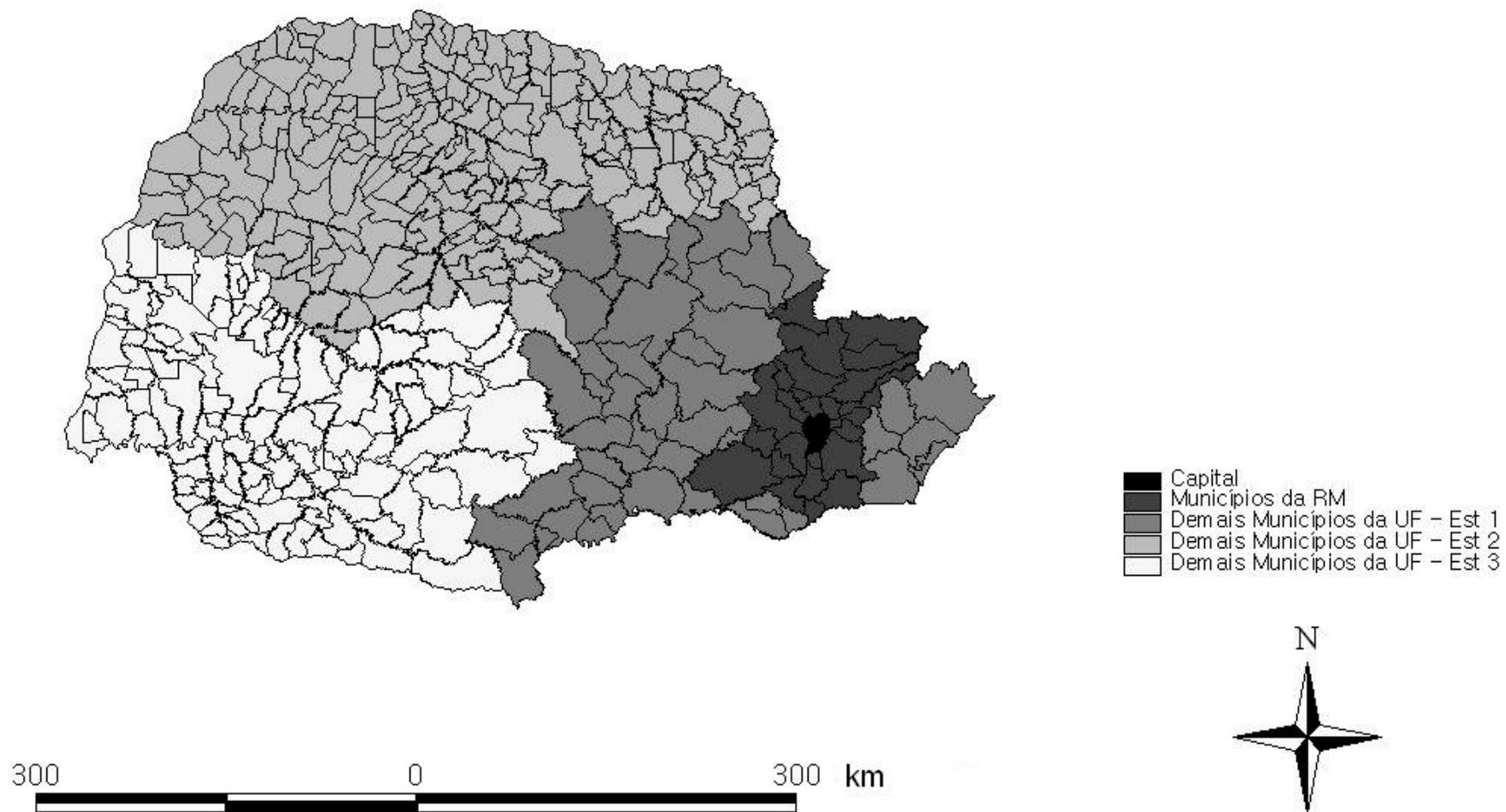
Mapa A1.19 - Distribuição dos municípios por estrato geográfico e espacial - Rio de Janeiro



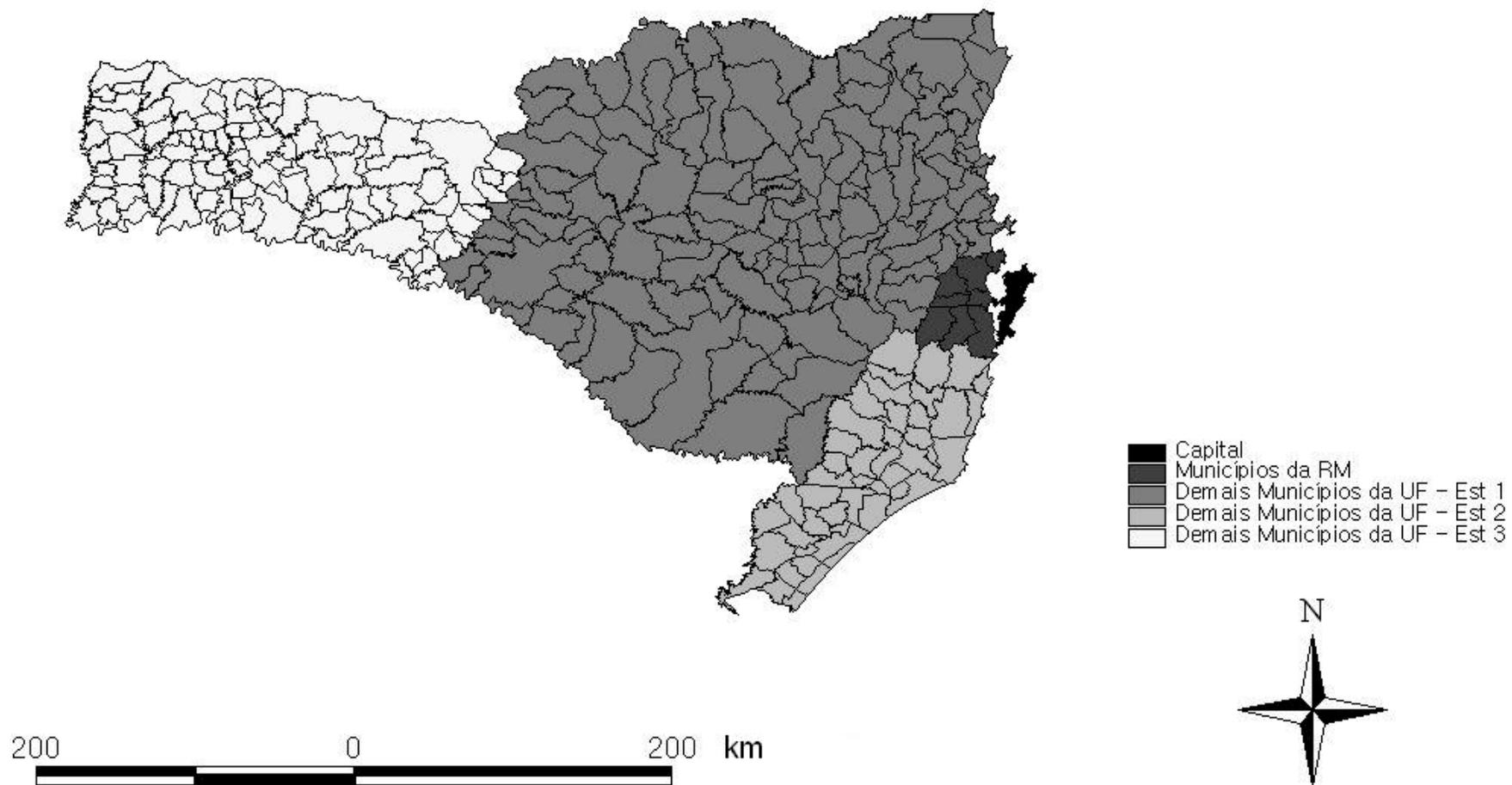
Mapa A1.20- Distribuição dos municípios por estrato geográfico e espacial - São Paulo



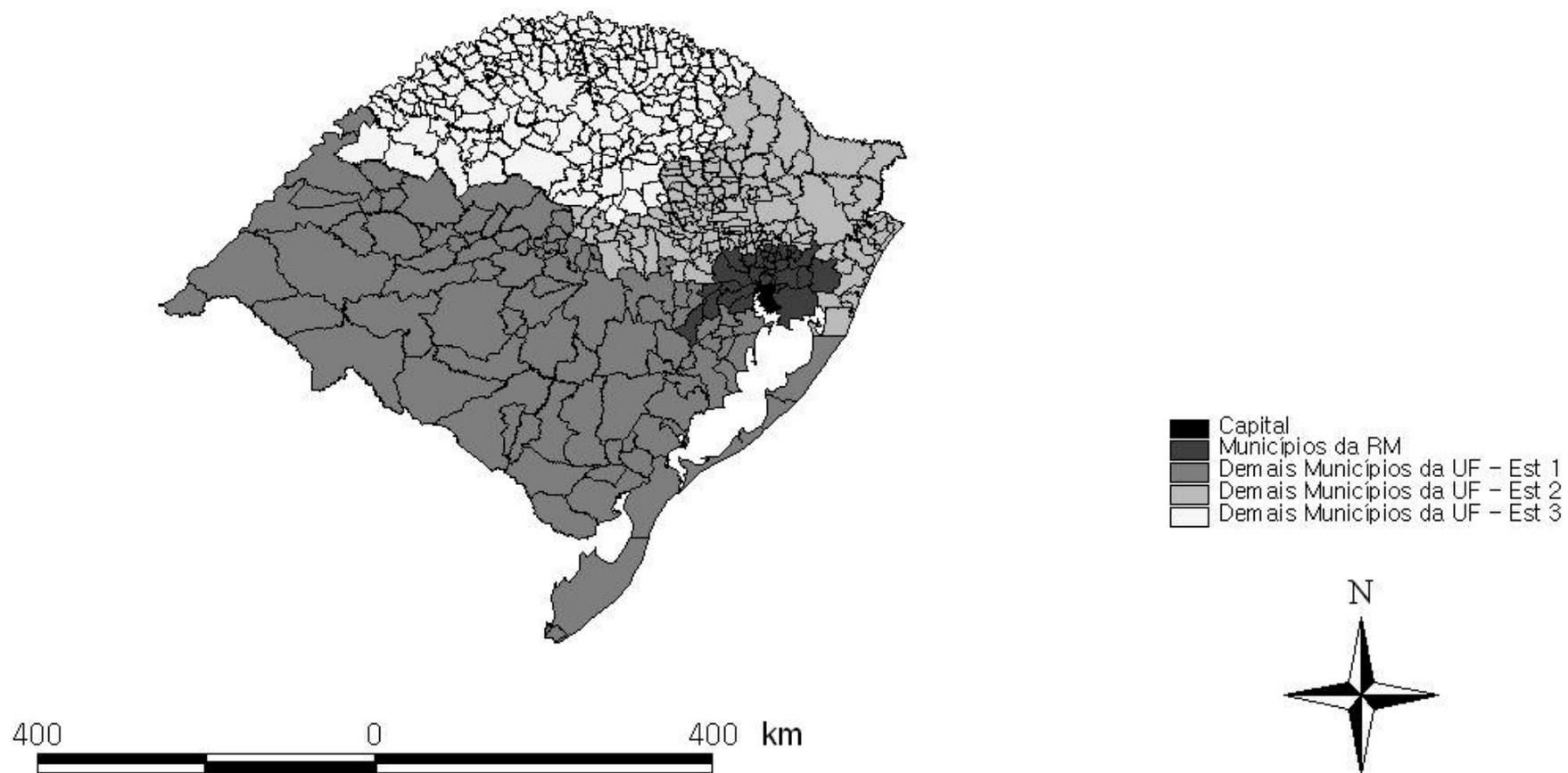
Mapa A1.21 - Distribuição dos municípios por estrato geográfico e espacial - Paraná



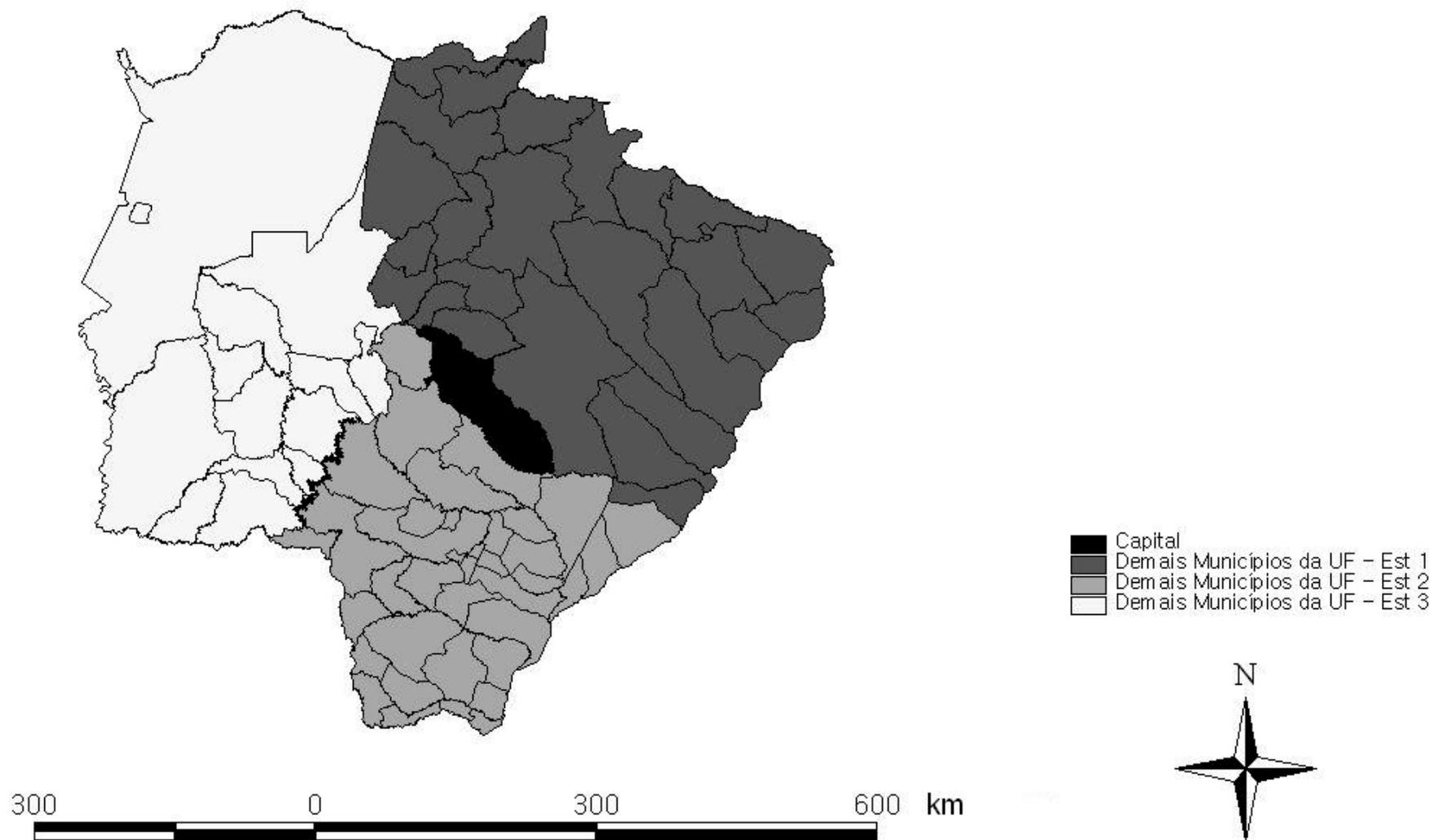
Mapa A1.22 - Distribuição dos municípios por estrato geográfico e espacial - Santa Catarina



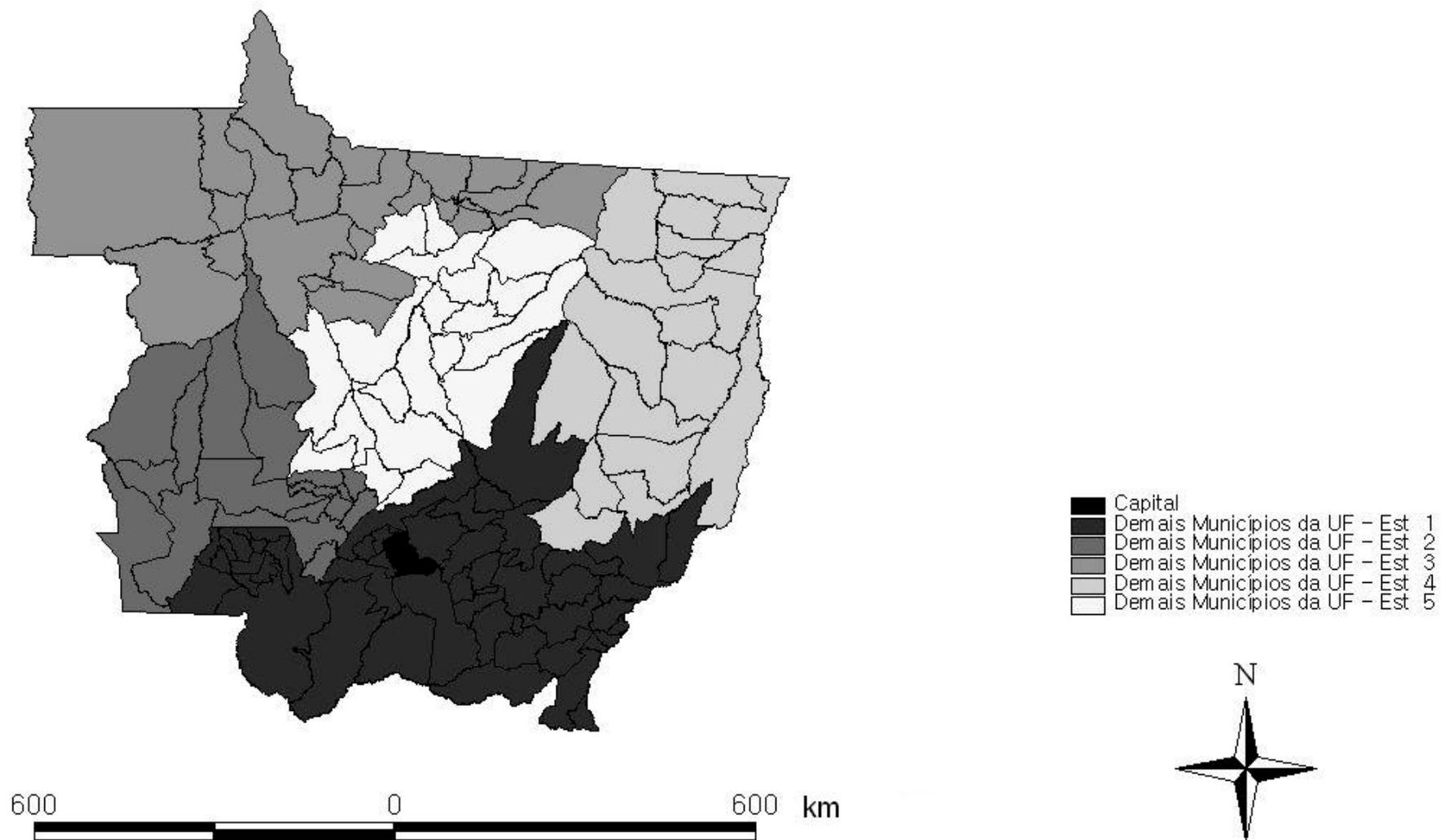
Mapa A1.23 - Distribuição dos municípios por estrato geográfico e espacial - Rio Grande do Sul



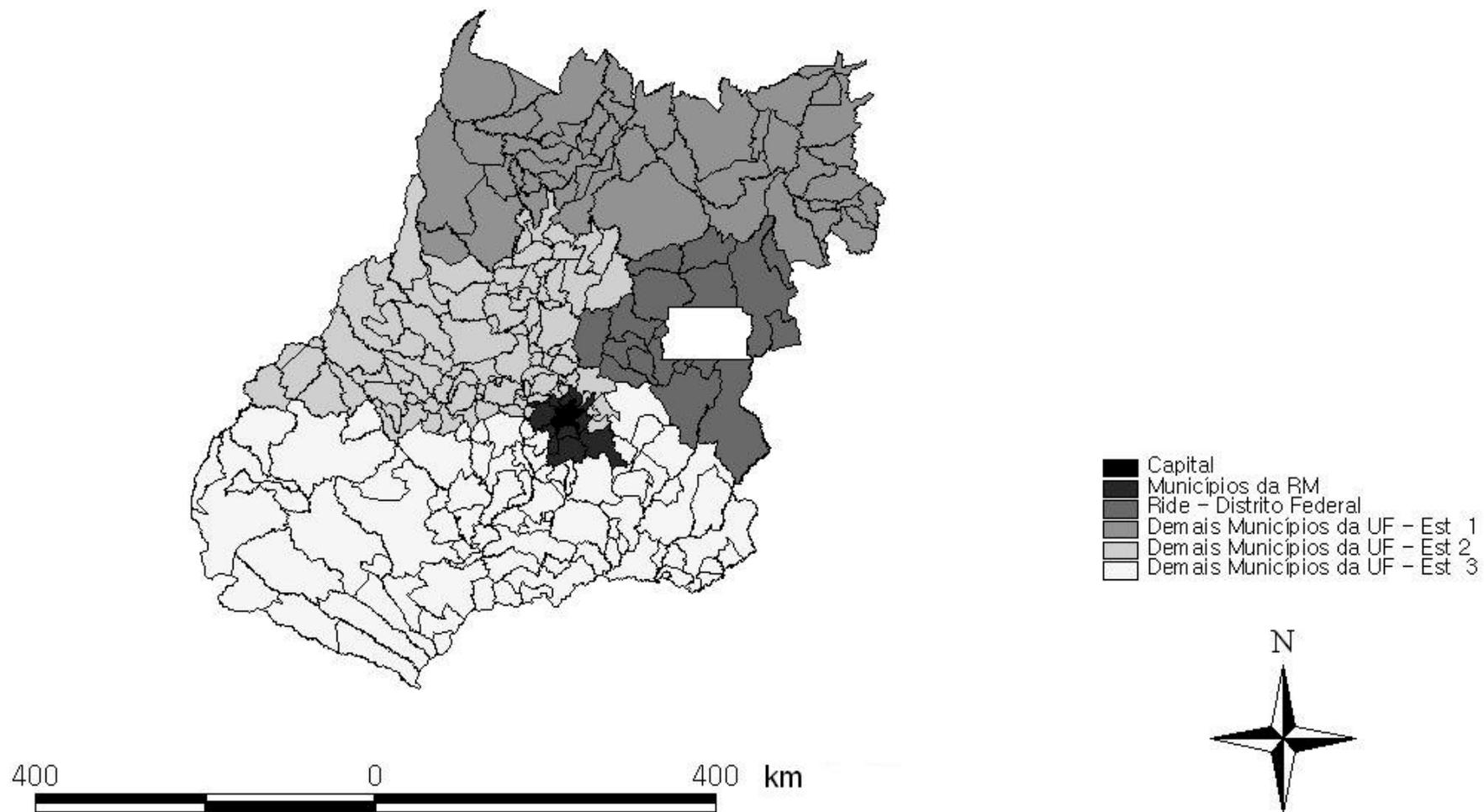
Mapa A1.24 - Distribuição dos municípios por estrato geográfico e espacial - Mato Grosso do Sul



Mapa A1.25 - Distribuição dos municípios por estrato geográfico e espacial - Mato Grosso



Mapa A1.26 - Distribuição dos municípios por estrato geográfico e espacial - Goiás



## Anexo 2 – Esquema de Rotação de Domicílios

Quadro A2.1 - Esquema de rotação: 1-2(5)

Mês	Painel																																							
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	C1	C2	C3	C4	C5	C6				
jan/09	5			4			3			2			1																											
fev/09		5			4			3			2			1																										
mar/09			5			4			3			2			1																									
abr/09				5			4			3			2			1																								
mai/09					5			4			3			2			1																							
jun/09						5			4			3			2			1																						
jul/09							5			4			3			2			1																					
ago/09								5			4			3			2			1																				
set/09									5			4			3			2			1																			
out/09										5			4			3			2			1																		
nov/09											5		4			3			2			1																		
dez/09												5		4			3			2			1																	
jan/10													5			4			3			2			1															
fev/10														5			4			3			2			1														
mar/10															5			4			3			2			1													
abr/10																	5			4			3			2			1											
mai/10																		5			4			3			2			1										
jun/10																			5			4			3			2			1									
jul/10																				5			4			3			2			1								
ago/10																					5			4			3			2			1							
set/10																						5			4			3			2			1						
out/10																							5			4			3			2			1					
nov/10																								5			4			3			2			1				
dez/10																									5			4			3			2			1			

## Anexo 3 - Tamanho da Amostra Mestra por domínios

Tabela A3.1 - Tamanho da amostra de setores para estimar o total de desocupados segundo as Regiões Metropolitanas, por ciclo de acumulação trimestral.

UF	Região Metropolitana ou RIDE	Total	Urbano	Rural
15 – PA	Belém	111	103	8
16 – AP	Amapá	54	45	9
17 – TO	Tocantins	54	46	8
21 – MA	São Luís	162	132	30
22 – RD	Teresina	120	92	28
23 – CE	Fortaleza	318	299	19
24 – RN	Natal	111	91	20
26 – PE	Recife	231	212	19
27 – AL	Maceió	215	207	8
29 – BA	Salvador	189	184	5
31 – MG	Belo Horizonte	231	212	19
32 – ES	Vitória	213	209	4
33 – RJ	Rio de Janeiro	453	445	8
35 – SP	São Paulo	450	332	118
41 – PR	Curitiba	216	178	38
42 – SC	Florianópolis	93	85	8
43 – RS	Porto Alegre	273	231	42
52 – GO	Goiânia	132	124	8
53 – DF	Distrito Federal	279	232	47

Tabela A3.2 - Tamanho da amostra de setores para estimar o total de desocupados por situação de setores segundo as capitais, por ciclo de acumulação trimestral

<b>UF</b>	<b>Capital</b>	<b>Total</b>	<b>Urbano</b>	<b>Rural</b>
11 – RO	Porto Velho	57	38	19
12 – AC	Rio Branco	51	46	5
13 – AM	Manaus	132	128	4
14 – RR	Boa Vista	51	47	4
15 – PA	Belém	69	65	4
16 – AP	Macapá	39	35	4
17 – TO	Palmas	39	35	4
21 – MA	São Luís	132	127	5
22 – PI	Teresina	69	65	4
23 – CE	Fortaleza	228	228	-
24 – RN	Natal	69	69	-
25 – PB	João Pessoa	60	60	-
26 – PE	Recife	96	96	-
27 – AL	Maceió	176	174	2
28 – SE	Aracaju	63	63	-
29 – BA	Salvador	147	146	1
31 – MG	B. Horizonte	129	129	-
32 – ES	Vitória	51	51	-
33 – RJ	Rio de Janeiro	234	234	-
35 – SP	São Paulo	273	193	80
41 – PR	Curitiba	117	117	-
42 – SC	Florianópolis	48	44	4
43 – RS	Porto Alegre	123	119	4
50 – MS	Campo Grande	75	71	4
51 – MT	Cuiabá	66	62	4
52 – GO	Goiânia	72	68	4
53 – DF	Distrito Federal	195	172	23