

# Relatório de projeto “Sistema Automático de Arrecadação” (Out/2012): estatística de leitura de placas.

---

Objetivo do projeto: Implantar pilotos do novo “*Sistema Automático de Arrecadação*” visando avaliar e testar as tecnologias disponíveis no mercado e, assim, determinar o sistema mais eficaz que atenda às normas e padrões de classificação veicular, de modo a estabelecer condições de contorno que permitam a otimização dos investimentos, utilizando os Recursos de Desenvolvimento Tecnológico (RDT), conforme o art. 2º da Resolução nº 483 da ANTT (Sistema Automático de Arrecadação para Teste, 2012).

## Sumário

Resumo .....	3
Introdução .....	5
Descrição e caracterização volumétrica dos sítios de aquisição .....	6
Apresentação da metodologia .....	11
Processo de aquisição .....	12
Discussão introdutória sobre as definições das notas de eficiência .....	15
Descritivo da tecnologia utilizada .....	18
Segmentação de dados e fundamentação estatística da análise utilizada.....	22
Sincronização de relógios .....	23
Processo de amostragem e fundamentação estatística .....	23
Tamanho das amostras usadas para análise de 17/9 e 18/9.....	27
Resumo da metodologia estatística utilizada .....	28
Representações analíticas das estimativas de notas (introdução).....	28
Cálculo do erro de uma nota e sua consolidação.....	30
Definições adicionais .....	31
Hierarquia binomial usada e notas finais.....	32
Correção das frações pelo fluxo real .....	36
Exemplo de cálculo de frações de classes de leitura (GRU 23/8/2012) .....	37
Processo de inspeção de imagens em BackOffice.....	40
Regras adicionais para inspeção das imagens.....	44
Análise de fluxo, velocidade e densidade de veículos .....	45
Descrição de resultados (análise de dados e tabelas com frações).....	46
Análise de dados (5/9/2012) .....	46
Frações descritas no relatório preliminar de 10/09/2012 .....	47

Análise de dados das 14:00 as 21:00 em 6/9, 7/9 e 8/9/2012. ....	49
Análise de dados (17/9/2012 e 18/9/2012): auditoria de tráfego.....	52
Tempos de inspeção em BackOffice .....	60
Análise de dados: evolução das frações (17/9/2012 e 18/9/2012) .....	62
Estatísticas específicas.....	74
Leitura automática de motos .....	74
Análise de um período com chuva (5:00 de 21/9/2012).....	74
Dimensão estimada de BackOffice .....	75
Índices de rastreabilidade de veículos potencial do sistema .....	79
Conclusões .....	82
Sugestões para estudos futuros .....	86
Apêndice 1 – Formatação de saída de dados para análise em BackOffice do projeto “Sistema Automático de Arrecadação para Teste” .....	88
Transmissão dos arquivos .....	88
Organização de Diretórios .....	89
1.1. Diretório Raiz.....	89
1.2. Diretório Diários .....	89
1.3. Subdiretório Pista .....	90
1.4. Subdiretórios de dados e imagens .....	90
1.5. Exemplo de construção de diretórios .....	91
Arquivo de Dados.....	95
Registro de Dados .....	95
Arquivo de Imagem.....	98
Identificação do fornecedor/local do pórtico .....	99
Data e Horário.....	100
1.6. Intervalos de horário.....	100
Registro de LOG.....	102
Apêndice 2 – Exemplo numérico de cálculo de índices e frações na estatística de inspeção de imagens. ....	103
Funções de dependência entre notas %ICE, %ICH e novas definições. ....	103
Subrotina para determinação dos índices e frações.....	104
Simulações numéricas .....	106
Apêndice 3 – Analise restrita de fluxo, velocidades instantâneas e densidade de tráfego para GRU.....	110

Bibliografia .....	120
--------------------	-----

## Resumo

Um sistema de teste de arrecadação foi instalado numa sequência de sítios ao longo da rodovia Presidente Dutra para a determinação do nível de acerto de inspeção automática via OCR para todos os períodos ao longo de um conjunto de dias de operação. A taxa efetiva de acerto (*considerada correção de perda de tráfego e apoio de BackOffice*, ou seja, é a soma %fCE+%fCH) para os melhores pontos flutua dentro do intervalo 70%-80%. Esse índice de validação de leituras foi obtido em apenas dois sítios, o de Guarulhos (km 227) e o de S. J. dos Campos ‘sul’ (km 140). O primeiro sítio adquiriu imagens da fronteira de veículos, enquanto que o segundo capturou imagens traseiras. Chamamos a atenção para o fato de que o GRU é responsável por 44% do fluxo de veículos no trecho de rodovia estudado pelo projeto, enquanto que o ponto SJS é responsável apenas por 12% do fluxo total. Para atingir o nível de certeza nos resultados, o projeto desenvolveu um processo de auditoria interna que levou em consideração a contagem de tráfego por meio de sistema independente (conhecidos como ATs). Considerando os dias ‘oficiais’ de aquisição, as taxas máximas de acerto corrigido para cada sítio se distribuem conforme segue:

- GRU: 71,5% (17/9/2012) e 72,7% (18/9/2012);
- SJS: 69,4% (17/9/2012) e 77,4% (18/9/2012);
- BMA: 22,0% (17/9/2012) e 23,0% (18/9/2012);
- SJM: 42,5% (17/9/2012) e 42,7% (18/9/2012);
- SJC: 26,4% (17/9/2012) e 24,7% (18/9/2012);

O projeto envolveu a integração de dados em diversos segmentos ou estratos tais como: sítio de aquisição, período do dia, tipo de dado dentre outros. Um processo especial de inspeção precedido por auditoria na estrutura de dados foi estabelecido. Notas preliminares (sem envolver estatística de tráfego) são descritas no relatório por completeza para conjunto de dias restritos (antes de 5/9 e conjunto 6/9, 7/9 e 8/9). Durante o conjunto de dados considerados ‘oficiais’, todos os fornecedores conseguiram operar seus sistemas de forma simultânea para possibilitar integralização de dados. O processo de inspeção final envolveu a utilização de 10 a 20 pessoas formando um comitê de inspetores (BackOffice) para agrupamento dos dados da leitura e detecção em várias classes: grupo de acerto do sistema automático, grupo de acerto exclusivo dos humanos, grupo de dados contendo placas ilegíveis, grupo de dados com placas inválidas e grupo de motocicletas. Cada um desses grupos foi usado para compilar índices ou frações de acerto específicas para cada estrato junto com seus desvios-padrão. O projeto demonstrou que existem variações no tempo de inspeção médio entre inspetores e que os pontos que apresentam melhor desempenho (essencialmente medido pela fração de imagens validadas do total de registros) também coincidem com os pontos de maior velocidade de análise em

BackOffice. Assim, quanto melhor for o desempenho do ponto de aquisição tanto menor é o esforço humano dispendido na análise das imagens. Isso acontece porque sistemas com índices de acerto automático piores requerem mais tempo de digitação de código correto e não apenas a mera checagem visual de imagens. Tempos típicos de análise em regime de alta eficiência estão no intervalo 2-4 segundos enquanto pontos que exigem digitação das placas fazem 8-12 segundos. O tempo médio de inspeção inferido a partir do ‘estado da arte’ instalado durante o projeto é da ordem de 4 segundos.

No que tange apenas às taxas de acerto automático, *sem considerar as perdas envolvidas*, para os melhores níveis alcançados (melhores sítios) não se observam variações grandes entre dia e noite, nem aparentemente com o tipo de placa (cor, reflexividade). Com consideração de análise de tráfego, foram observadas perdas consideráveis de registro nos dias 17/9 e 18/9 para os sítios BMA e SJC (índices de relação registro-tráfego superiores a 60%). Para esses pontos, as notas de inspeção automática foram corrigidas para o fluxo registrado e resultaram em valores típicos entre 20%-30%, bem inferiores aos registros de GRU e SJS. O sítio de SJM, embora tenha apresentado o menor índice de perda de veículos, apresentou o menor nível de acerto de todos os sítios. De forma geral, as perdas por ilegibilidade de placa ou imagens não conforme somam aproximadamente entre 10%-20% do fluxo efetivamente registrado para os sítios com as melhores notas de inspeção, enquanto que o resíduo das diferenças para o total deve necessariamente ser inspecionado por operadores humanos. O esforço de BackOffice (percentagem de imagens do total que deve necessariamente ser inspecionado por humanos) envolve considerações financeiras e pode ser estimado parcialmente pela análise dos índices. Porém, com base no tempo médio determinado para inspeção de um registro pelo BackOffice usado no projeto, foi estimado o número de pessoas necessárias para inspecionar *todos os registros* gerados por todos os pontos, assumindo que *não há perdas de tráfego* (registro). Para o tempo médio de inspeção (4s), 48 pessoas são necessárias trabalhando em um turno de 8h, enquanto que se o tempo de 10s de digitação média for usado (o operador não faz a checagem do código gerado, mas digita diretamente a placa) esse mesmo número é da ordem de 120 pessoas. Variações de duração de turno podem ser importantes para garantir boa eficiência de inspeção pelos humanos. Adiantamos também que os erros de leitura automática observados (ou, equivalentemente, frações de imagens exclusivamente lidas por humanos) estão muito relacionados à presença de caracteres específicos que causam maior erro. Tais caracteres podem ser usados para segregar placas em BackOffice, aumentando a eficiência final de inspeção. Ainda na questão ‘detecção de veículos’, ficou claro que o uso de laços virtuais – em que pese a vantagem expressiva de não se ter que instalar nenhum dispositivo na pista – sofre de algumas desvantagens, como a necessidade de regime de iluminação constante.

Com relação à incidência de chuvas, boa parte do projeto transcorreu durante o período de estiagem. Por causa disso, uma análise sistemática da influência de chuvas não foi realizada, o que exigiria, rigorosamente, o uso de ferramentas de medição pluviométricas em cada sítio. Entretanto, no dia 21/9/2012 foi possível extrair as frações de desempenho para um período de chuva relativamente moderada para GRU, o que foi observado por meio da impressão da imagem (perda de contraste). Os resultados não permitem inferir que houve perda de qualidade no processo de inspeção, utilizando dados de um dos melhores pontos de aquisição, embora tenha ocorrido um aumento pequeno da taxa de imagens ilegíveis. Com relação à eficiência de inspeção de motocicletas, um dos sítios (SJS) capturou imagens traseiras, e valores típicos da ordem de 25% foram encontrados. A maior parte dos registros de motos (aproximadamente 50%) somente pode ser inspecionada por operadores humanos.

No que se refere à possibilidade de se usar o sistema para rastreamento sucessivo de veículos, ou, de forma equivalente, as taxas de leitura automática sucessivas, a utilização das melhores frações de validação de código e sob a assunção de perda nula de tráfego levam a valores teóricos da ordem de  $81,9\% \pm 9,0\%$  para 2 pórticos,  $72,9\% \pm 12,1\%$  para 3 pórticos e  $65,6\% \pm 14,5\%$  para uma sequência de 4 pórticos. Durante o projeto não foi possível executar um teste de rastreabilidade, o que não impediu a determinação dos coeficientes de rastreamento, que podem ser facilmente obtidos a partir dessas taxas de leitura individuais.

## Introdução

Em vista da crescente necessidade de suporte ao volume de veículos em circulação pela malha rodoviária nacional e, com vistas a melhor atender o usuário dentro de um serviço continuado de alto-padrão, o domínio das técnicas automáticas de identificação e classificação de veículos se torna ferramenta indispensável.

A dinâmica de circulação de veículos nos principais eixos de transporte do país se caracteriza pela complexidade de cenários de densidade e tipos de veículos. Nesse contexto, fica evidente a necessidade de se criar, ao mesmo tempo, condições para aumentar o fluxo combinado com uma maior capacidade de identificação e classificação dos veículos que circulam pelo sistema.

Essa necessidade combinada é um desafio técnico e operacional, já que identificação e classificação são em princípio mais difíceis de alcançar quando os veículos estão em movimento, em alta velocidade e nas variadas condições de trânsito comuns nas Rodovias Brasileiras.

Um sistema de identificação e classificação de veículos automático operacional nestas condições é uma ferramenta útil para estabelecer o conhecimento do tipo de usuário das Rodovias – origem / destino, legalidade das condições dos veículos e dinâmica de causa-e-efeito no tráfego pelos grandes eixos. A partir do

estabelecimento dessas condições de identificação e classificação automáticas, será possível estruturar uma operação de pedágio mais justo, que considera trechos percorridos, tipos de veículo e condições de trânsito.

O projeto teve como objetivo montar, comissionar, integrar dados e gerar estatísticas de índices de acerto (Silva, et al., 2006; Evaluation of the license plate recognition system, 2012) provenientes de 5 pórticos localizados na rodovia Dutra sobre a responsabilidade da concessionária ‘Nova Dutra’. Cada um dos pórticos possui infraestrutura específica para inspeção de veículos fornecida por empresas diferentes, o que fornece variedade de tecnologias ao projeto, embora, essencialmente, os métodos de leitura e recursos de aquisição ópticos sejam os mesmos. Assim, seria esperado que variações observadas nos resultados reflitam pouco a variedade de tecnologias que participam do projeto e mais as condições extrínsecas ou ambientais. Por exemplo, espera-se que a taxa de imagens absolutamente não conformes<sup>1</sup> oscile mais como função da densidade de veículos sob o sistema de aquisição e não pelo fato de se usar um tipo de câmera ou outro. A taxa de imagens não conforme é um dos resultados que se extraem dos dados mais críticos, uma vez que placas que não podem ser lidas absolutamente representam uma perda de cobrança no modelo (Silva, et al., 2006) que prevê cobrança de pedágio por leitura de imagem de placa<sup>2</sup>.

Por outro lado, a taxa de imagens conformes para um humano é diretamente relacionada à chance de, em uma passagem, o sistema globalmente resultar em um registro que só pode ser lido por um agente humano. Essa frequência multiplicada pelo fluxo de veículos é de importância fundamental para o dimensionamento do BackOffice ou procedimento de validação de imagens não lidas automaticamente.

## Descrição e caracterização volumétrica dos sítios de aquisição

O objetivo fundamental da fase preliminar do projeto foi medir a qualidade do processo de leitura automática de placas de veículos através de um conjunto de 5 portais localizados na rodovia Dutra, administrados pela concessionária Nova Dutra<sup>3</sup>, na sua maior parte dispostos no sentido Rio de Janeiro → São Paulo. A posição dos 5 pórticos é como segue (ver Figura 3):

1. Km 227 SP (Guarulhos);
2. Km 140 (S. J. dos Campos);
3. Km 266 (Barra Mansa/RJ);
4. Km 167 (S. João do Meriti/RJ);

---

<sup>1</sup> Isto é, imagens que não podem ser lidas nem por um software OCR (software de leitura automática) e nem por um humano.

<sup>2</sup> Também chamado de ‘video tolling’ em inglês

<sup>3</sup> <http://www.novadutra.com.br/>



5. Km 156 (São José dos Campos/SP): esse pórtico adicional ao conjunto original de 4 pórticos, captura dados no sentido São Paulo → Rio;



Figura 1 Imagens que ilustram alguns pontos originais de registro de dados de placa que fazem parte do projeto, no sentido horário da esquerda para a direita: km 227, km 266, km 167 e km 156.

Imagens de aspectos a partir da rodovia de cada um dos pórticos (exceto pelo 5º sítio na pista norte) podem ser vistas na Figura 1. É importante considerar a variedade de número de pistas envolvido no estudo. Por exemplo, o pórtico do km 227 tem cerca de 8 pistas (incluindo acostamento). Já o pórtico no km 226 tem apenas 2 pistas e 1 acostamento. Essa condição diversificada de pistas representa uma variabilidade de condições de aquisição, configurações possíveis e fatores extrínsecos de interferência externa. A Figura 2 traz o mapa da distribuição geográfica da rodovia que abrigou os sistemas de aquisição.

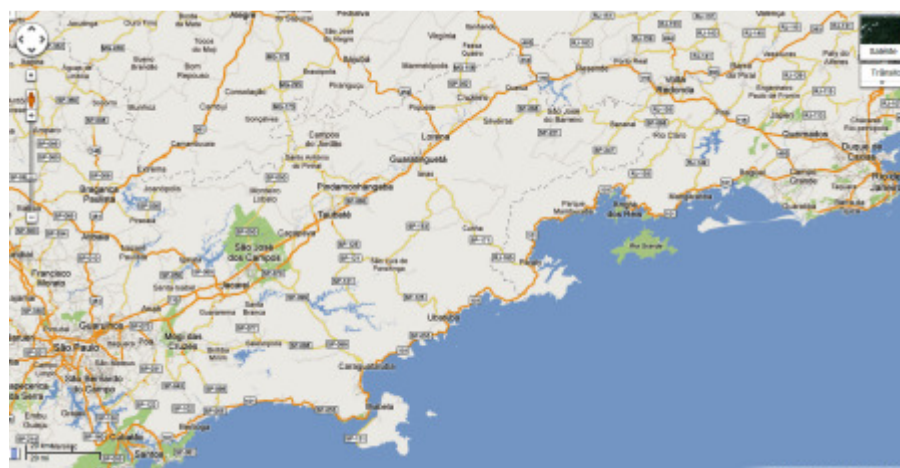


Figura 2 Trecho de rodovia considerado no projeto. Fonte: Google.

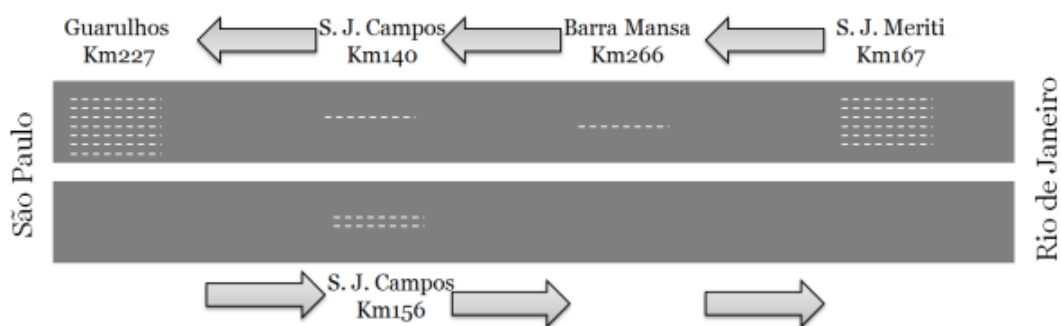


Figura 3 Estrutura de faixas e sequência da posição dos sítios de leitura (pórticos). A pista superior é a pista ‘sul’.

Assim, cada sítio é caracterizado por um número diferente de faixas ou pistas, bem como volumes (fluxos) diferenciados. A Tabela 1 mostra a integração de fluxos obtidos em 17/9 e 18/9 durante o projeto com base em dados de contadores de tráfego que foram instalados no local.

	17/set	18/set	17/set	18/set	17/set	18/set	17/set	18/set	17/set	18/set
	GRU		SJS		BMA		SJM		SJC	
P1	17861	18034	24269	24382	8265	8368	7382	7385	22847	22803
P2	15744	16107	17508	17883	12065	12282	7690	7517	18416	18865
P3	23628	23526					9	109	10124	10447
P4	26310	26010					24854	27441		
P5	25845	25303					22005	23789		
P6	24131	24030					12496	11015		
P7	14138	14630					1741	1036		
P8	1404	1856								
Subtotais	149061	149496	41777	42265	20330	20650	76177	78292	51387	52115
Totais por sítio	298557		84042		40980		154469		103502	
Médias diárias	149278,5		42021		20490		77234,5		51751	
Pesos por sítio	43,8%		12,3%		6,0%		22,7%		15,2%	
Total	681550									

Tabela 1 Fluxos integrados de tráfego em 17/9 e 18/9 por sítio e por pista.



Nessa tabela é fornecido o fluxo integrado para os dias, por sítio e por pista, o que permite calcular os volumes médios diários, os totais por sítio, os pesos por sítio e o total de fluxo nesses dois dias.

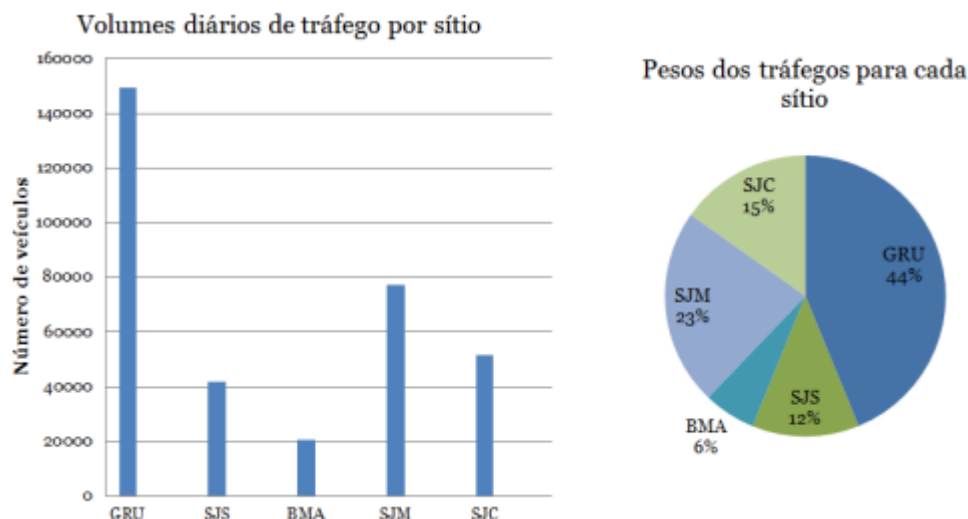


Figura 4 volumes diários (esquerda) e pesos de participação de volumes por sítio.

Para facilitar a interpretação desses dados de fluxo, a Figura 4 traz um histograma dos fluxos médios estimados por sítio de aferição e os pesos de participação de cada sítio no projeto de acordo com esse volume. É importante considerar que o sítio GRU é responsável por 44% do fluxo total, o que equivale aos fluxos de SJC, SJM e BMA juntos. Essa observação é importante não só para a escolha dos pontos onde se deve instalar o sistema de cobrança, mas para se ponderar os resultados deste relatório conforme o grau de esforço que foi dispendido em cada sítio.

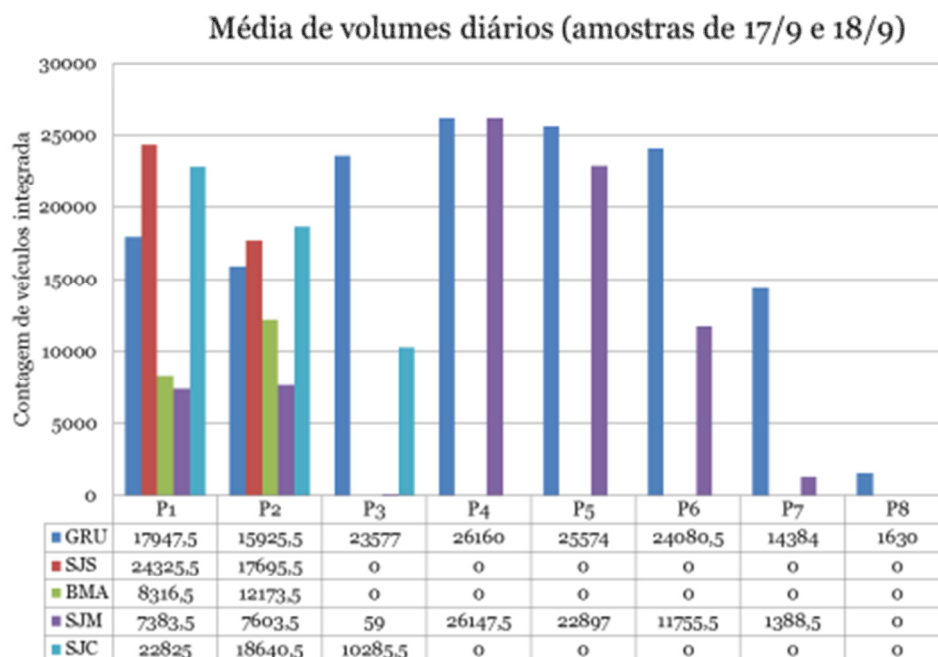


Figura 5 Distribuição de fluxos horários integrados por sítio e pista usando amostras de 17/9/2012 e 18/9/2012.

A distribuição de fluxos diários por pista pode ser visto na Figura 5. Essa figura representa a média por pista dos fluxos integrados para os dias 17/9 e 18/9. O percentual de ocupação do fluxo diário por pista para cada pista pode ser visto na Tabela 2. Como se pode ver, a ocupação das pistas é razoavelmente homogênea para todos os sítios com uma queda prevista para as pistas mais lentas ou mais rápidas (por causa da velocidade). Esse percentual mostra que é necessário cobrir todos os sistemas de pista, mas que sazonalmente equipamentos podem ser ‘desligados’ conforme a ocupação para redução de custos de manutenção e operação em pistas pouco ocupadas.

	Pesos por faixa				
	GRU	SJS	BMA	SJM	SJC
P1	12,0%	57,9%	40,6%	9,6%	44,1%
P2	10,7%	42,1%	59,4%	9,8%	36,0%
P3	15,8%			0,1%	19,9%
P4	17,5%			33,9%	
P5	17,1%			29,6%	
P6	16,1%			15,2%	
P7	9,6%			1,8%	
P8	1,1%				

Tabela 2 Percentual de ocupação das pistas por sítio.

## Apresentação da metodologia

A metodologia utilizada foi essencialmente estatística (Xavier Jr., 2012). A Figura 6 é uma descrição esquemática sumária do processo de inspeção utilizado que precede a análise estatística. Em tempo real é feita a aquisição dos dados (em cada um dos pórticos). A aquisição fornece os dados pertinentes que podem ser divididos nos tipos:

1. *Referências de posição de aquisição e tempo*: informação que fornece o momento da aquisição, origem do dado (por exemplo, qual foi o subsistema responsável pela aquisição) e resultado de leitura automática;
2. *Dados de passagem*: informações adicionais relacionadas a sensores como de velocidade e comprimento;
3. *Imagem*: conteúdo 2D que fornece a entrada para inspeção. Esse conteúdo deve também ser único e correspondente ao registro de passagem. Deve permitir análise posterior em BackOffice do conteúdo (releitura).

A formatação dos dados obedece à estruturação que está descrita no Apêndice 1. Toda a infraestrutura e segurança dos dados, mecanismos de transferência e informações de contexto para prevenção de fraudes, foram providos e são de inteira responsabilidade da empresa Engelog-CCR. Essencialmente, é preciso garantir:

- A consistência dos dados (se houve parada de aquisição, qual a causa etc);
- Que cada passagem seja única: mais de uma passagem registrada para um mesmo veículo implica em aumento não justificado de esforço de BackOffice;
- Que sejam coincidentes com o momento de sua geração: isso é o que descrevemos como ‘tempo de latência nulo’, não pode haver intervalo de tempo entre a aquisição e a contabilização dos dados para evitar fraudes;
- Que haja informação sobre a origem dos dados (de que sistema ele procede): de novo, isso tem a ver com a necessidade de se manter a consistência dos dados;
- Que se garanta a presença de apenas uma imagem para cada passagem registrada: excesso de imagens para um mesmo registro (passagem) gera gargalos desnecessários no tempo de transferência de informação para BackOffice,
- Que as imagens tenham qualidade mínima de leitura tanto para o processo automático como para os humanos: é preciso que uma ‘prova’ suficiente seja apresentada para justificar o processo de leitura. Essa ‘prova’ deve ser legível pelo humano.

Além disso, um conjunto de informações adicionais que podem ou não ser capturadas no momento da passagem são fornecidas. Essas incluem informação sobre velocidade

e comprimento do veículo. Relatórios de log do sistema<sup>4</sup> foram requisitados para fornecer o contexto sistêmico no momento de passagem.

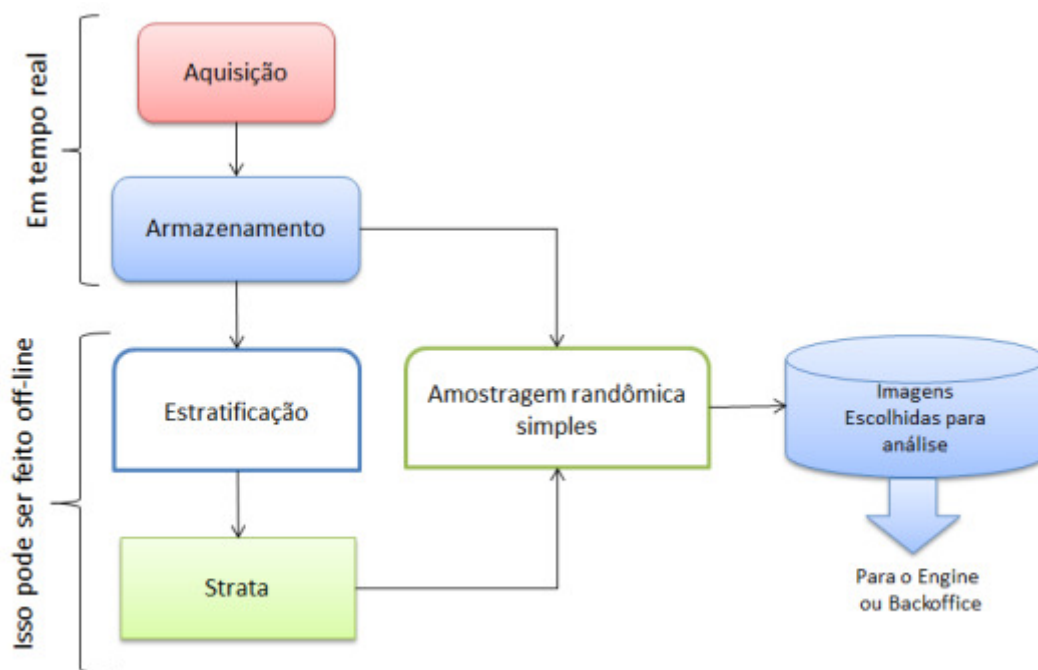


Figura 6 Diagrama de blocos representativos do processo de inspeção que foi implementado durante o projeto.

### Processo de aquisição

O processo de aquisição só termina quando as informações são armazenadas de forma conveniente conforme ilustra o fluxo da Figura 6. Uma vez armazenadas em um dispositivo específico, as informações são adequadamente separadas para submissão à análise estatística. Isso envolve um conjunto de procedimentos que podem ou não se valer de informação externa adicional e pertinente ao momento em que as informações foram colhidas (por isso, o momento de registro ou passagem deve ser fornecido).

O processo de aquisição (Figura 6) é seguido de uma estratificação inicial de dados onde ‘*strata*’ são gerados, o que permite que um processo de amostragem randômica seja aplicado de forma eficiente. Os dados escolhidos são então encaminhados para BackOffice (inspeção pelos humanos).

Em termos específicos, as informações que podem ser utilizadas para a análise (chamamos de estratificação, ver Figura 6) podem ser divididas nos seguintes tipos:

1. Data (formato: Ano/mês/dia/hora/minuto/segundo);
2. Condições de tráfego (livre, laminar, congestionado: depende da velocidade do tráfego);

<sup>4</sup> Esses relatórios devem conter dados como: momentos de início de aquisição, descritivo de ocorrência de erros, paradas, descrições de status de aquisição etc.

3. Identificação da câmera que realizou a aquisição;
4. Dados climáticos e de insolação;
5. Dados de posição do sol (azimute e elevação do sol);
6. Status da iluminação (on/off);

Essas informações são, porém, inter-relacionadas. Por exemplo, a posição do sol em relação à placa (que é importante para se determinar o grau de iluminação diurno) pode ser obtida a partir da data e da geometria de posicionamento da câmera. A identificação da câmera é importante para fornecer informações sobre a posição da fonte de dado. Não se procedeu à aquisição de dados climáticos e de insolação<sup>5</sup> em conjunto com a aquisição de dados.

☞ Durante o período de aquisição realizado, apenas a data e origem dos dados foram utilizadas. A qualidade da imagem foi avaliada diretamente por inspeção humana e nenhuma correlação climática foi feita. Na ocorrência de modificações grandes na qualidade das imagens, procedeu-se à coleta de dados independente de fontes de dados externas.

É importante considerar que, mantidos inalterados os parâmetros do software OCR, não há diferença de resultado se a leitura de imagem for feita em tempo real ou não. O resultado da leitura pelo software OCR não é um atributo da passagem ou da imagem (softwares OCR diferentes podem retornar leituras diferentes para a mesma imagem), razão porque fornecedores diferentes foram usados.

É possível também considerar uma lista dos fatores que afetam a qualidade da imagem no momento da passagem. Tais fatores tem impacto decisivo na nota de classificação final e, principalmente, na ‘taxa de acerto do software OCR’. Há fatores, porém, que são de grande importância na estimativa da limitação do método de inspeção, por exemplo, a qualidade intrínseca das placas em análise. Essa importância justifica uma segmentação de dados específica para isolar essa característica.

1. Resolução da imagem;
  - a. Dimensão dos pixels no sensor: dependente do tipo de câmera utilizada;
  - b. Objetiva (parte óptica): depende do arranjo óptico utilizado e da distância (altura) da câmera em relação à placa;
2. Iluminação;

---

<sup>5</sup> Dados de insolação exigem a instalação de sensores específicos nos pontos de captura. Dadas as distâncias envolvidas (extensões de 500km), todos os pontos deveriam contar com sensores de insolação e climáticos. Por outro lado, o período muito curto de aquisição dos dados não coincidiu com o da incidência de chuvas, tornando o monitoramento pluviométrico desnecessário.

- a. Tempo de exposição: parâmetro configurável que depende da potência do sistema de iluminação;
- b. Iluminação externa;
  - I. Altura e posicionamento do sol;
  - II. Insolação (quantidade de nuvens);
  - III. Potência do iluminador e status (on/off);

- c. Contraste: depende do nível de iluminação externo (não só devido ao sol, mas ao céu e presença de objetos que causem sombreamento, por exemplo, árvores, prédios);



Figura 7 Tipos de placa em circulação no Brasil.

- 3. Posicionamento da placa na imagem;

- a. Altura da placa em relação ao ponto de trigger: depende de detalhes de instalação (e. g., distância do ponto de trigger do pórtico, ângulo de posição da câmera, altura do pórtico etc);

- b. Velocidade do veículo: a velocidade afeta diretamente o tempo de ‘permanência’ do veículo na região de leitura e interage fortemente com o tempo de exposição (quanto maior o tempo de exposição tanto menor é o limite de velocidade para o qual se garante uma boa taxa de leitura);

- c. Variedade entre placa dianteira e traseira;

- 4. Tipo de placa;

- a. Cor: no caso do Brasil, há placas com vários tipos de cor de fundo (Figura 7), por exemplo, fundo vermelho (caminhões) (DENATRAN/CGIE, 2006). A cor está relacionada com o contraste final da imagem e é um parâmetro potencialmente importante que justificaria segmentação de dados específica<sup>6</sup>;

<sup>6</sup> Ou seja, fornecer notas de desempenho específicas para variedades de placas com cores diferentes.



- b. Estado de conservação: efeito de particular importância ligado diretamente à frequência de placas ilegíveis por nenhum método e que representa perda de arrecadação;
5. Tipo de letras presentes na placa: é fácil compreender que a maior parte dos erros de leitura, tanto do sistema automático como de humanos ocorre com algumas letras que têm formatos parecidos. Estatísticas especiais para essas ocorrências podem ser determinadas.

### **Discussão introdutória sobre as definições das notas de eficiência**

Um ‘nota’ é um número que representará o status de saída ou resultado de inspeção para uma determinada base de dados de entrada. Compreensivelmente uma imagem deve conter informação suficiente para se extrair um código. Portanto os elementos fundamentais da análise são:

1. Registro: conteúdo de informação que identifica uma passagem associada a uma imagem;
2. Imagem: conteúdo de informação que pode conter uma placa;
3. Placa: pedaço da imagem que pode conter informação sobre um código;
4. Código: sequência de caracteres que pode representar um código válido;
5. Código válido: pedaço de informação que pode conter um código validado (placa correta) do veículo capturado na imagem;
6. Código validado: informação que contém o código existente e correspondente ao do veículo responsável pelo registro.

O fluxo de informação da placa até o resultado final desejado é Registro→Imagem→Placa→Código→ Código válido→ Código validado (ver também seção “Dimensão estimada de BackOffice” e Figura 73). A chance de se obter uma leitura correta de um veículo registrado no momento da passagem é o produto das probabilidades:

- Probabilidade de se capturar um veículo: essa probabilidade depende de fatores sistêmicos como, por exemplo, eficiência de disparo da câmera no momento da passagem do veículo;
- Probabilidade de se capturar uma placa: dado que uma imagem foi adquirida contendo um veículo, ela deve conter uma placa. Essa probabilidade depende de um grande número de fatores extrínsecos e intrínsecos ao sistema: qualidade da placa, ângulo de posição do par câmera-veículo, iluminação externa, presença de sobras, ângulo de posição do sol etc;
- Probabilidade de se ter um código válido: uma vez capturada uma imagem contendo uma placa, uma fração da imagem será lida com o algoritmo OCR e resultará em um código que pode ser válido. Por código válido entende-se uma sequência de 3 caracteres de letras e 4 caracteres

de números. Essa probabilidade depende da qualidade da imagem e da eficiência de detecção da placa na imagem pelo software;

- Probabilidade de se ter uma leitura validada: depende de fatores de BackOffice (habilidade do inspetor) de validar a imagem, ângulo de posicionamento, iluminação e tipos de caracteres presentes na placa.

A chance de se ter uma leitura correta (validada) será, então, o produto das probabilidades descritas acima. Não é difícil ver que, se um código válido existe no registro, então, muito provavelmente, uma placa legível por um humano será encontrada na imagem. Isso somente não ocorre se, por má fé, códigos ‘chutados’ foram inseridos propositalmente nos registros. Mas, nesse caso, a segregação de imagens conforme os códigos válidos levará a detecção desses ‘chutes’.

A eficiência de leitura pode ser avaliada diretamente a partir da inspeção realizada por humanos sobre as placas. Por exemplo, ao se examinar por meio de um agente humano, um grupo de  $n$  imagens (amostra) contendo, cada uma, placas únicas de veículos escolhidos em determinada hora de um determinado dia, o índice de acerto para o humano (ou nota de imagens conformes para o humano, %ICH<sup>7</sup>) pode ser definido como:

%ICH = número de imagens legíveis apenas pelos humanos/Número total de imagens ilegíveis para software OCR,

Esse é o número de imagens *incorretamente* interpretadas pelo software OCR, mas que, para um operador humano bem treinado, resultam na leitura correta dos caracteres. Uma vez que se espera que cada placa contenha sete caracteres, é possível também especificar %ICH<sub>2</sub> (definição 2) como função do número total de caracteres

%ICH<sub>2</sub> = número de caracteres lidos na amostra  $n/(7*n)$ ,

É possível ver que %ICH ≤ %ICH<sub>2</sub> e, portanto, determinar %ICH é necessário e suficiente<sup>8</sup>. Da mesma forma, podemos definir:

%ICE: nota de imagens legíveis pelo software OCR. Trata-se da frequência de acertos do software (ou de conjunto de softwares) e envolve, necessariamente, a análise por BackOffice via operadores humanos. Podemos também definir essa nota por: %ICE = número de imagens corretamente lidas pelo software OCR/N;

%INCE: nota de imagens não lidas pelos softwares de OCR. Trata-se de %ICE descontado de 100%, ou seja, é o resíduo das leituras conformes o software OCR;

---

<sup>7</sup> Observamos que, o número de imagens lidas por um humano pode ser muito maior que simplesmente  $n*%ICH$ , pois um humano pode ser obrigado a checar se o software OCR resultou em uma leitura correta, ou seja, verificar todas as ICE. O que determina essa obrigação é uma decisão de investimento e não uma característica do sistema.

<sup>8</sup> Não basta que o mesmo conteúdo de caracteres de letras e números estejam contido no código, é necessário que estejam também na ordem correta!

%INC: nota de imagens absolutamente ilegíveis. É o resíduo das imagens ilegíveis pelos humanos (%ICH).

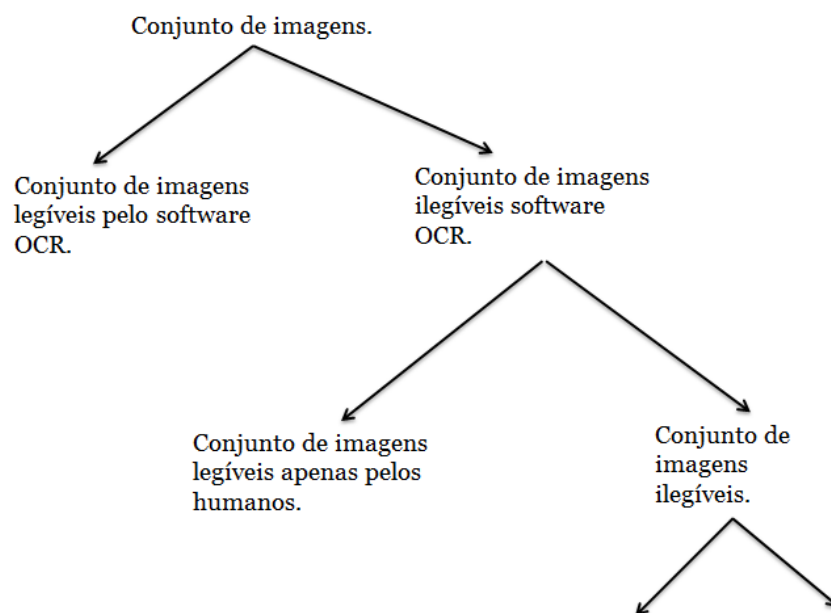


Figura 8 Diagrama de uma das alternativas para análise de um conjunto de  $n$  imagens para apresentação do problema a ser tratado. Abaixo do nível das imagens ilegíveis, há outras separações.

Como a eficiência de leitura de um humano é idealmente superior ao de softwares OCR, %INC é o percentual esperado de imagens ilegíveis. As causas de perda de reconhecimento são:

1. Imagens com pouca iluminação (por motivos vários, chuva, neve etc);
2. Presença de caracteres problemáticos;
3. Placas avariadas irreconhecíveis;
4. Ocultamento devido a excesso de veículos (efeito de sombra causado por veículos altos e densidade elevada de tráfego);
5. Falhas técnicas de disparo de imagem (causas ligadas ao sistema de captura);

Esses fatores têm impacto direto na probabilidade de se ter um código válido, uma vez capturada uma imagem contendo uma placa. De qualquer forma, as características acima descritas podem ser inferidas a partir do conjunto de imagens. Portanto, o %INC pode ser subdividido em diversos outros índices menores. Em particular, os fatores 2 e 3 são os mais importantes pois:

- O fator ‘placas avariadas irreconhecíveis’ determinam diretrizes para substituição de placas;
- Caracteres (letras) graficamente parecidos, o que causam erros de reconhecimento tanto pelo sistema automático como pelos humanos;

Em menor grau, o fator ‘ocultamento devido ao excesso de veículos’ representa uma perda dependente do nível de densidade de veículos (veículos por km), portanto, se

essa taxa de inconformidade fora do valor tolerável, pórticos de registro de placas para fins de pedágio devem eventualmente se localizar em regiões livres de congestionamento (ou regiões onde congestionamento sejam menos frequentes). Tais questões devem ser respondidas dentro do escopo do projeto.

A Figura 8 representa um diagrama esquemático de um relacionamentos entre as notas que é concebível para um conjunto de imagens. Não é difícil ver que são válidas as relações:

$$\%ICE + \%INCE = 1$$

$$\%ICH + \%INC = 1$$

Relações desse tipo podem ser obtidas para vários outros índices como é feito na Seção “Hierarquia binomial usada e notas finais”. Além disso, é importante considerar que essas notas são sazonais, isto é, elas dependem do tempo. Isso acontece principalmente por causa de variações na densidade de veículos e variações no nível de iluminação externa. Por causa dessa dependência, é razoável montar o processo de amostragem de dados *por período*. Idealmente não é necessário definir um ‘mínimo período’, mas amostrar sobre um período em que se observem a ocorrência de condições externas homogêneas, ou seja, o processo de amostragem estatística deve ser feito sobre fatias homogêneas de dados. Mais informação sobre os índices utilizados no estudo, consultar a Seção “Hierarquia binomial usada e notas finais”.

## Descritivo da tecnologia utilizada

Nesta seção descrevemos brevemente a tecnologia que foi majoritariamente utilizada no estudo e que envolveu esforço coordenado de várias empresas diferentes. Sistemas de visão de máquina ou sistemas de visão automatizada (N. Sebe, 2005) têm sido largamente utilizados na indústria para aumentar a qualidade de itens manufaturados, e a aplicação direta de sistemas específicos de reconhecimento de padrões (bem como aumento da velocidade de processadores) possibilitou desenvolver dispositivos que fazem isso em tempo real. Em particular, sistemas de leitura de caracteres (Grant, 2012?; Ozbay & Ercelebi, 2005) tiveram sucesso na automatização do processo de transcrição de caracteres representados de forma escrita para o formato digital. A representação digital possibilita a universalização da representação de dados, além de permitir armazenamento por tempo indeterminado, facilitando ainda a tarefa de transmissão (Hadjipanayis, 2009).

Não demorou muito até que sistemas de visão de máquina fossem adaptados para inspeção de placas de veículos. A partir da tecnologia desenvolvida para OCR, sistemas de ALPR<sup>9</sup> foram desenvolvidos, ou seja, os mesmos sistemas de

---

<sup>9</sup> Do inglês: *automatic license plate recognition*.

representação e armazenamento de dados foram modificados para cumprir a tarefa de inspeção de placas de veículos em tempo real.

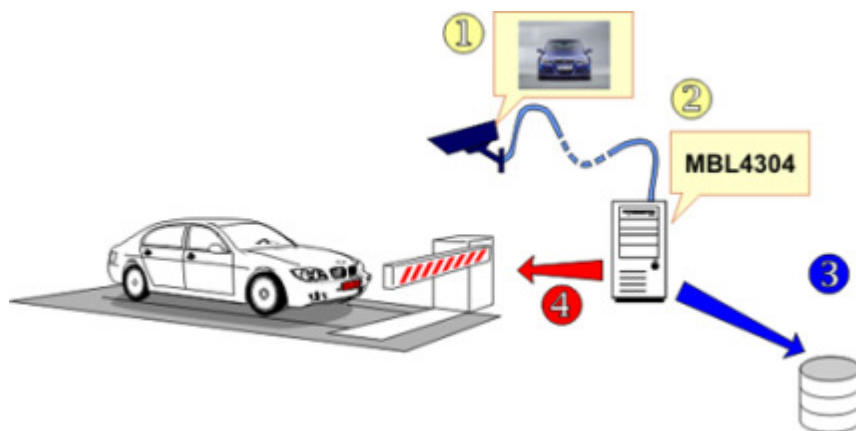


Figura 9 Elementos essenciais de um sistema de registro e leitura automático de placas de veículos.

Um sistema de registro automático de placas é composto dos elementos:

1. Câmera (Figura 9-1) composta por elemento sensível, óptica, sistema de disparo e iluminação artificial;
2. Sistema de registro de passagem: essencialmente um PC ou processador (Figura 9-2) capaz de receber a imagem digital ou digitalizada e executar o procedimento de leitura automática (por meio de um software OCR);
3. Banco de dados para armazenamento de resultados (Figura 9-3);
4. Opcionalmente, a partir do resultado da inspeção realizada em tempo real, pode-se obter um sinal que libere um dispositivo em pista, por exemplo, uma cancela (Figura 9-4).

Complementarmente, um ‘escritório auxiliar’ (*BackOffice*) pode ser responsável por validar os resultados de leituras gerados pelo processo automático. O *BackOffice* é essencialmente um comitê de humanos treinados e reconhecidamente mais capazes do que máquinas na inspeção de caracteres contidos nas imagens de placas de veículos.



Figura 10 Segmentação de um caractere na imagem.



Figura 11 Interface API típica para verificação e programação de LPR.

A primeira tarefa do software de leitura de placas é determinar se há uma placa na imagem. Se isso for verdade, ela tem que ser localizada corretamente e, então, os caracteres devem ser processados. Depois de localizados na imagem, os caracteres são transformados em um formato padronizado o que implica em: normalização das cores do fundo (background via normalização de contraste e brilho), normalização de tamanho e orientação da placa.

O próximo passo é a *segmentação dos caracteres* (Figura 10): depois de extrair a região de presença dos caracteres, eles devem ser distinguidos (segmentados) uns dos outros. A segmentação torna-se uma tarefa difícil quando a imagem da placa não é clara suficiente (presença de ruído, configuração deficiente de atributos do sistema de aquisição, como ajuste de foco e uso de lentes inapropriadas). Isso ocorre quando os caracteres se misturam ou outros efeitos de luz (presença de sombra) estão presentes na placa. Uma vez que os caracteres são segmentados (separados uns dos outros), o algoritmo de reconhecimento de imagem é invocado para interpretar cada segmento que foi separado na imagem. A resposta desse processo é o código ASCII correspondente ao caractere em particular.

O processo de aquisição na câmera pode ser único ou múltiplo. No primeiro caso, apenas uma imagem é registrada para inspeção automática. No segundo caso, mais de uma imagem é registrada e submetida ao processo de inspeção. O processo de inspeção automático múltiplo pode ainda fazer uso de regras específicas de identificação. Por exemplo, na maior parte das placas no Brasil, espera-se uma sequência de 3 letras do alfabeto seguida de 4 números conforme as regras do RENAVAM. Qualquer outra combinação que resulte da leitura automática poderá ser



descartada (Draghici, 1997). Isso implica em um nível de adaptação da tecnologia (pós-processamento) a convenção gráfica das placas usadas localmente. Por outro lado, no nível do sistema de leitura também é possível estabelecer filtros para aumentar o grau de confiabilidade da leitura resultante. Por exemplo, se existe confusão por parte do algoritmo entre a letra ‘T’ e o número ‘1’, é possível interpretar todos os ‘T’ do 2º campo (numérico) eventualmente lidos como ‘1’ e eventuais ‘1’s lidos no 1º campo como ‘T’s. O mesmo vale para o ‘O’ e o ‘o’ e assim por diante. Sistemas de identificação de placa para fins de multa por infração de trânsito já dispõem desse recurso de pós-processamento.

Para ser útil, o sistema de leitura automático de placas deve contar ainda com uma API<sup>10</sup>. Uma API é uma série de funções que torna a tarefa de programação e ajuste no sistema de inspeção menos difícil através de uma interface entre o processo de inspeção e as imagens brutas, a fim de que programadores humanos possam ajustar e melhorar a eficiência de leitura.

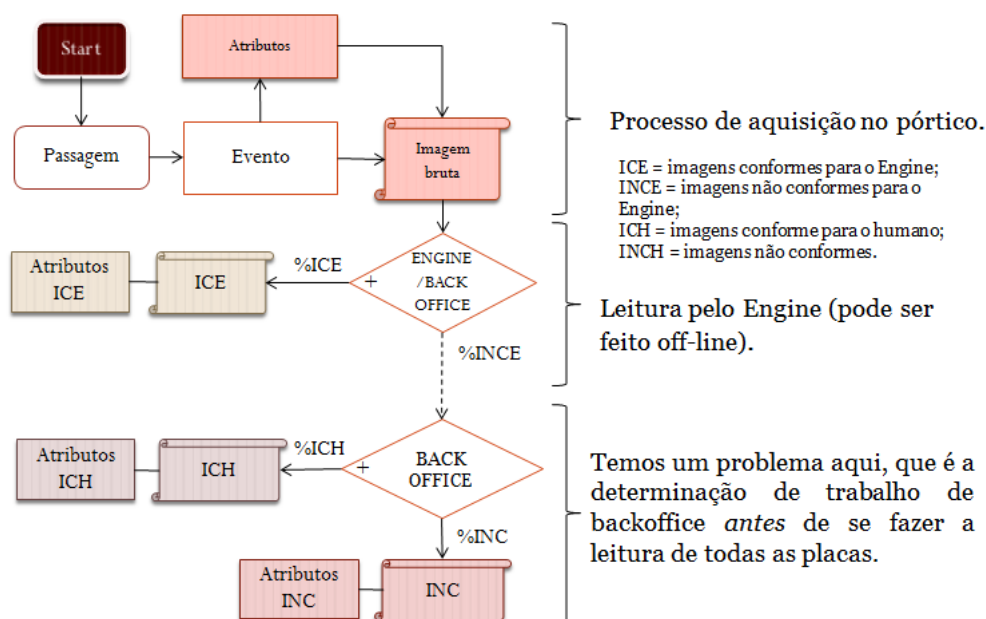


Figura 12 Diagrama em blocos do processo de aquisição e inspeção.

É importante também considerar brevemente o sistema que gera a captura da imagem. A câmera pode funcionar em modo ‘intermitente’ (registro de vídeo em movimento) e um processo de inspeção quadro a quadro pode ser utilizado. O mais comum é uma câmera sobre uma pista controlada por um dispositivo externo e independente conhecido como ‘trigger’ (disparo). Esse é responsável por disparar a câmera na ocorrência da passagem de um veículo. A partir do disparo, a câmera pode adquirir uma ou várias tomadas. Modernamente, câmeras evoluíram de forma a dispensar a instalação de sistemas de disparo na pista. Isso teve como motivação a redução da complexidade de instalação de sistemas de LPR, dispensando

<sup>10</sup> Do inglês: *Application Programming Interface*.

intervenções de parada de pista<sup>11</sup>. Sistemas de ‘trigger virtual’ permitem que se adquiram imagens ou grupo de imagens quando uma determinada seção do vídeo é invadida por uma alteração de brilho (o que ocorre com a passagem do veículo). Há, entretanto, o inconveniente de sempre se precisar de iluminação externa durante períodos de pouca luz (noite). O conjunto de sistemas instalados no projeto apresenta uma mistura entre diversos tipos de tecnologias quanto ao processo de controle das câmeras. O efeito do uso de um ou outro tipo de sistema de disparo, se detectável, surge na qualidade da imagem, erros de disponibilização de registros (perda de passagem) ou na posição da placa em relação ao frame da imagem. Entretanto, em sistemas tecnologicamente maduros, variações detectáveis pelo uso de um (trigger físico) ou de outro (trigger virtual) sistema não são relevantes. Entretanto, os resultados apresentados neste relatório mostram claramente a ocorrência dessa variação *inter* sítios por causa das várias tecnologias utilizadas.

## **Segmentação de dados e fundamentação estatística da análise utilizada.**

O processo de aquisição em cada um dos portais que fazem parte do projeto está descrito de forma esquemática na Figura 12. Idealmente uma vez iniciado o processo de aquisição no pórtico, um evento é gerado a partir de uma passagem. O evento é caracterizado por um conjunto de atributos associados a uma imagem. O conjunto formado pela imagem e atributos (registro) é submetido então ao processo de inspeção automatizado (software OCR). Em uma situação ideal (em que se sabe que uma determinada imagem foi corretamente interpretada pelo software OCR), dois grupos de diretórios-destino devem ser preenchidos: o grupo dos diretórios que contém imagens corretamente interpretadas (ICE) e o grupo de imagens incorretas (INCE). No estágio de consolidação de dados, todas as imagens devem ser inspecionadas por humanos, o que fornece as taxas de acerto do software OCR em uma primeira abordagem. As imagens incorretas<sup>12</sup> devem então ser submetidas ao processo de inspeção humana (BackOffice) para se interpretar corretamente aquelas que, não obstante serem conformes para um humano, não puderam ser interpretadas de forma automática. Portanto, o resultado final da inspeção humana também prevê dois diretórios-destino: o das imagens corretamente interpretadas por humano (ICH) e o das imagens absolutamente não conformes. Posteriormente, as imagens não conformes podem ser subclassificadas conforme a causa da inconformidade (imagem inválida, placa ilegível etc). O processo de análise descrito dessa forma é apenas um dos inúmeros processos que se podem estabelecer.

---

<sup>11</sup> Essas intervenções seriam necessárias no momento da instalação da estrutura de sustentação das câmeras e com a manutenção periódica dos sistemas de pista.

<sup>12</sup> Ou seja, imagens reconhecidamente lidas incorretamente pelo sistema automático, mas que podem conter informação legível por um humano.

Entretanto, para que o estudo presente seja viável, não será possível analisar *todas as imagens* que são produzidas, dados os fluxos envolvidos. Então, nesta seção discutimos processo de amostragem que permitem inferir corretamente as taxas de acerto (índices) e o intervalo de validade.

### Sincronização de relógios

O processo de captura e registro de informação deve ser feito de forma que a informação temporal seja sincronizada para todos os pórticos. Isso é garantido fazendo-se com que o relógio de cada pórtico opere segundo um relógio único (não é necessário sincronização com horário oficial) que é naturalmente o relógio do sistema de integração de dados. Por esse meio, garante-se que os dados coletados em regiões tão diferentes possam ser integrados horário a horário, garantindo-se observação de variação de índice de acerto (p. ex., fluxo de placas ilegíveis) tanto globais como locais.

### Processo de amostragem e fundamentação estatística

Amostragem é um processo de seleção de dados de um conjunto usualmente extenso chamado população (Spiegel, 1975). Por meio da amostragem, estimam-se parâmetros estatísticos dessa população sem que se tenha que interagir com todos os elementos que a compõem. Nesse sentido algumas questões válidas são: quais são os melhores ‘estimadores’ de parâmetros estatísticos? Para esses estimadores, qual é o erro associado? De forma geral, para uma população com  $N$  indivíduos, uma amostra envolve a escolha de  $n$  ( $n < N$ ) elementos. Bons estimadores para os parâmetros característicos  $P$  e  $A$  são tais que:

$$\lim_{n \rightarrow N} p(n) = P,$$

$$\lim_{n \rightarrow N} a(n) = A,$$

Ou, de outra forma, o valor ideal para  $n$  deve obedecer às relações:

$$|P - p(n)| < \varepsilon_p,$$

$$|A - a(n)| < \varepsilon_a,$$

onde  $\varepsilon_p$  e  $\varepsilon_a$  são números positivos e menores que 1 ligados ao erro máximo esperado na estimativa.

O processo de amostragem deve sempre proceder à escolha de elementos da população de forma aleatória. Portanto, do ponto de vista prático, cada elemento da população acessado em certa ordem pode ser escolhido através de um número aleatório gerado no momento da escolha. A partir disso é possível programar dois tipos de procedimentos para amostragem:

- *Amostragem com substituição*: quando o elemento escolhido é devolvido à população e uma nova escolha é feita;

- *Amostragem sem substituição*: o elemento é retirado da amostra. Com isso a população é reduzida numa próxima iteração de amostragem;

Extraí-se um conjunto de  $n$  elementos  $X$  de uma população com  $N$  indivíduos. Dentre os estimadores típicos que são usados para caracterizar uma população, os mais clássicos são:

$$E(X) = \mu_x = \mu \quad (\text{média}),$$

A teoria estatística garante que a média estimada com base na amostra é um bom estimador para a média da população. *Se a amostragem se der com substituição*, então a variância da estimativa das médias da amostra será dada por:

$$E[(X - \mu)^2] = \sigma_x^2 = \frac{\sigma^2}{n},$$

onde  $\sigma$  é a variância da população. Se a amostragem não for com substituição, então a variância da estimativa das médias será dada por:

$$\sigma_x^2 = \frac{\sigma^2}{n} \left( \frac{N - n}{N - 1} \right).$$

Portanto, se a população for muito grande, não há diferença entre as estimativas das variâncias das médias com ou sem substituição. Esse é o caso do problema de inspeção de placas.

O problema da inferência dos índices de acerto de leitura de placas é semelhante ao problema de se selecionar um elemento de certo tipo em uma população de  $N$  elementos. Deve-se determinar qual a quantidade aproximada de elementos do tipo particular, conhecendo-se o total de elementos e executando um processo de amostragem. Ao se retirar um elemento da população, esta deve ser devolvida a ela para que a amostragem ocorra ‘com substituição’. Repete-se o processo  $n$  vezes e uma estimativa do total de elementos de certo tipo será dado pelo valor médio do número de elementos extraídas com o tipo correspondente dividido por  $n$ . No processo de análise de imagens temos, portanto:

- A população com  $N$  elementos é o total de imagens em certo período de tempo (estratificação de dados), a ‘cor’ é o resultado da análise do BackOffice para um índice pretendido, por exemplo:
  - Inspeção do software OCR OK;
  - Inspeção pelo software OCR incorreta<sup>13</sup>;
- Por outro lado, se a imagem não for lida pelo software OCR ela pode ser:
  - Lida por um humano;
  - Não lida por um humano (na situação de placa avariada ou ausente da imagem).

---

<sup>13</sup> Para que isso ocorra todos os caracteres da placa devem ser corretamente lidos.

Há, portanto, uma hierarquia de ‘problemas de elementos’ para os quais se aplica a *estatística da distribuição binomial* (problema da urna ou lançamento da moeda<sup>14</sup>). A estimativa da média é, portanto, o índice de acerto de inspeção e a variância estimada fornece uma medida do erro dessa estimativa. Observa-se, além disso, que o resultado de uma inspeção é independente do outro. Uma vez que o problema pode ser descrito pela distribuição binomial, então se a taxa ou índice esperado é  $t$  então:

$$\mu_t = t,$$

e

$$\sigma_t = \sqrt{\frac{t(1-t)}{n}}$$

é o desvio padrão esperado para a estimativa das médias. Em toda estatística, portanto, o resultado será fornecido conforme a relação:

$$\mu \pm \sqrt{\frac{t(1-t)}{n}},$$

*É importante considerar que, para  $n > 30$ , o intervalo fornecido é a taxa medida em que 68,2% das ocorrências amostradas são observadas ( $\pm 1$  sigma). Ou seja, se o experimento for realizado 100 vezes, em aproximadamente 68 repetições a taxa medida cairá no intervalo inferido. É possível também considerar a notação de intervalo:*

$$\left[ \mu - \sqrt{\frac{t(1-t)}{n}}, \mu + \sqrt{\frac{t(1-t)}{n}} \right].$$

Uma estimativa numérica obtida a partir de simulação pode ser vista na Figura 13. Essa figura representa um estudo numérico de desvio da estimativa da média de uma amostra com valores diferentes de  $t$ . No eixo ‘y’ representam-se esses desvios para  $t = 50\%$ ,  $65\%$ ,  $80\%$  e  $99\%$  usando um conjunto de 30 amostras. O valor ‘N’ na Figura 13 corresponde ao ‘n’ do texto e é o número de elementos em cada amostra.

<sup>14</sup> No caso de lançamento de moeda  $t=50\%$ .

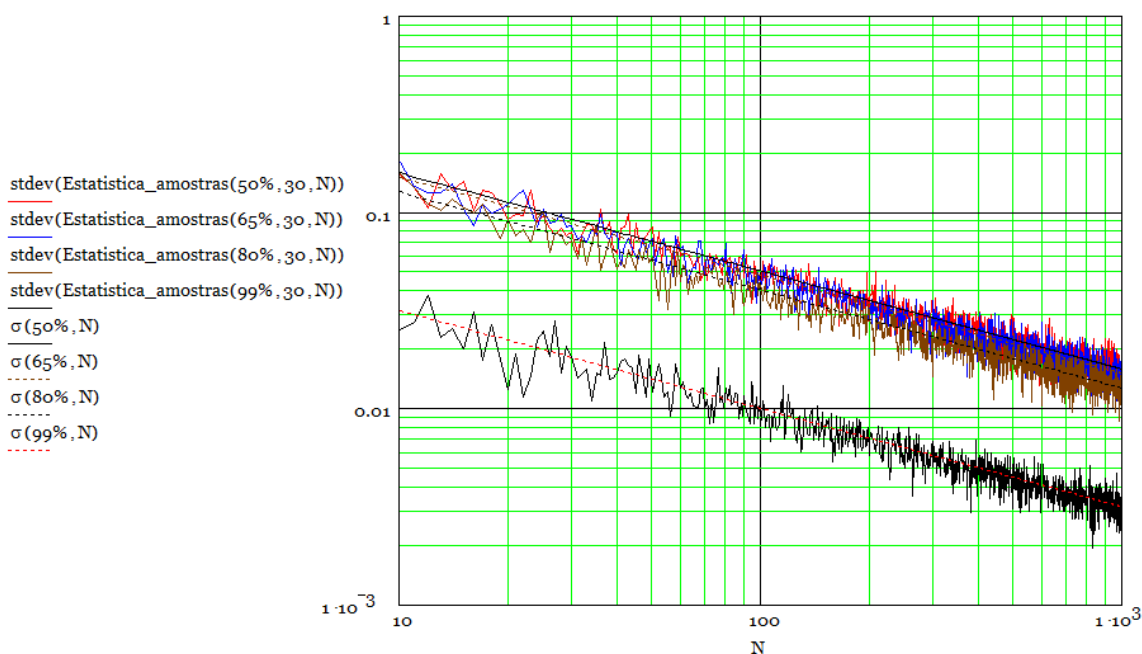


Figura 13 Comportamento do desvio padrão com função do número de elementos extraídos de amostra simulada superposta à curva teórica de acordo com uma distribuição binomial (Erro ou desvio das estimativas de médias).

A Figura 13 mostra que, a partir de 20 elementos, o erro da estimativa da média é inferior a 10% pp<sup>15</sup>. Para taxas próximas a 50%, com 100 elementos, o nível de erro é de 5%pp. Observe que há uma dependência do nível de erro com a taxa próximo a 0% ou 100% de taxa, o número de elementos de uma amostra necessário para que o erro esteja abaixo de certo valor é consideravelmente menor (aproximação para a certeza absoluta, ver Figura 14). Para que o erro seja da ordem 1%pp, com notas próximas a 50%, o número de elementos da amostra tem que ser maior que 1000.

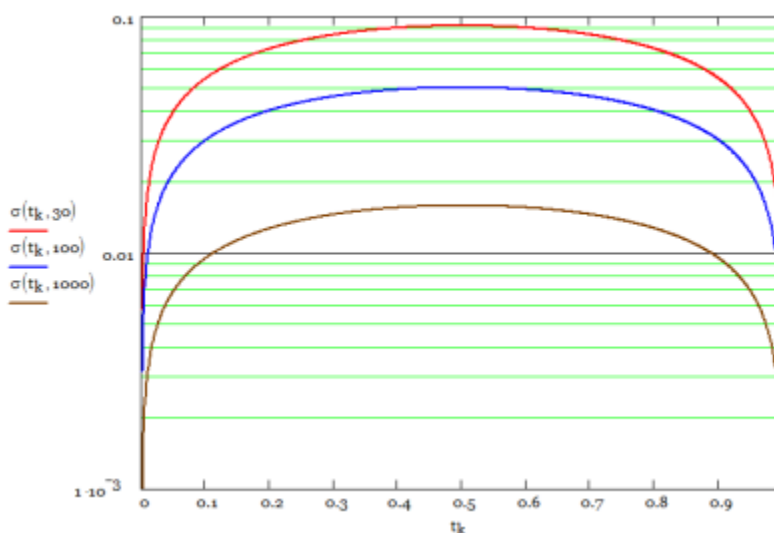


Figura 14 Dependência do erro da estimativa da média com a taxa.

<sup>15</sup> pp = pontos percentuais. Note que isso é diferente de %, ou seja, 5%pp de 60% determina um intervalo de variação de [55%,65%], enquanto que 5% de 60% é compreensivelmente igual a 3,0%.



Com o objetivo de tornar o estudo exequível, optou-se por utilizar  $n \leq 100$  para cada extrato pré-definido nos períodos de maior variabilidade de condições externas. O fato de o número de elementos da amostra ser fixo impõe regularidade sobre o processo de amostragem, impedindo que períodos diferentes tenham número de elementos diferentes. O intervalo de erro dependerá, portanto, da taxa média de acerto (índice) característico do conjunto de dados em exame, sendo limitado a 5%pp para taxas em torno de 50%. A redução do intervalo é obtida integrando-se dados de vários dias consecutivos para os quais se observou as mesmas condições externas.

### **Tamanho das amostras usadas para análise de 17/9 e 18/9.**

De acordo com a distribuição de sítios de coleta definidos no projeto, há uma estratificação automática por ‘sítios de leitura’. Portanto, é possível retornar índices de acerto para o quesito de leitura de placa para cada sítio que representa uma demonstração espacial da estatística.

Na representação dos dados coletados, o menor intervalo de integração de dados<sup>16</sup> é 30 minutos. Nesse intervalo de tempo, um arquivo de captura contém uma quantidade variável de passagens (registros) associadas idealmente a imagens que devem conter placas legíveis. Portanto, o processo de amostragem fica automaticamente estratificado em intervalos de 30 minutos. No âmbito temporal, a estratificação pode ser feita aplicando-se o processo de amostragem para outros intervalos pré-definidos. Entretanto, uma vez que não se espera observar variações de fatores externos durante o período noturno, é possível, por exemplo, estratificar os dados no intervalo 19:00 a 05:00 (do dia seguinte) em intervalos de maior duração. Com isso evita-se a realização redundante de inspeção na faixa de horário noturno, o que não contribui para a estatística por conta da pouca variabilidade das condições externas. A partir da estratificação temporal, é possível determinar os momentos de melhor ajuste das notas e, com isso, obter correlação com fatores externos<sup>17</sup>.

Um resumo do número máximo de elementos de amostragem por estrato temporal para os períodos considerados ‘oficiais’ (dias 17/9 e 18/9) pode ser visto na Tabela 3. Essencialmente, esses valores foram utilizados para que houvesse um mínimo de representatividade estatística por período de forma e, ao mesmo tempo, não comprometer a data de entrega do relatório que permaneceu fixa, enquanto que o cronograma do projeto apresentou atrasos por conta de ajustes nos sistemas de aquisição de imagens em vários pontos. Em outras palavras, esses atrasos reduziram consideravelmente o tempo disponível para inspeção de BackOffice, o que definiu o tamanho das amostras. Essencialmente, o número de registros por sítio levou em conta o tamanho do sítio em termos de número de pistas, a quantidade de períodos básicos (96) e o tempo disponível para a inspeção. Um conjunto de 11 inspetores foi utilizado. Com base em análises preliminares de tempo de inspeção por registro (da ordem de 5 segundos), um tempo máximo de inspeção foi calculado em 35,8 horas. É

<sup>16</sup> Esse é o intervalo com que os dados são armazenados pelo conjunto de câmeras em um arquivo individual e despachados para o servidor central. Portanto, durante 24 horas, há 48 arquivos gerados.

<sup>17</sup> Dentro os quais se destacam variações por conta de mudanças de iluminação (posição do sol).

preciso não se esquecer de que, além da inspeção das imagens, os dados devem ser consolidados, integrados e manipulados, o que exigiu outro tempo, já que o processo de consolidação não está automatizado.

	Períodos	Pistas	Registro/período	total Registros
<b>BMA</b>	96	2	<b>20</b>	3.840
<b>GRU</b>	96	8	<b>10</b>	7.680
<b>SJC</b>	96	3	<b>13</b>	3.744
<b>SJM</b>	96	7	<b>10</b>	6.720
<b>SJS</b>	96	2	<b>20</b>	3.840

Tabela 3 Amostragem usada com o período simultâneo de aquisições.

Na demonstração dos resultados, dada a disponibilidade de dados de aferição de tráfego apenas em intervalos horários, *as frações foram exibidas em intervalos horários assim como os dados de tráfego*, que levou a integração dos períodos básicos para obedecer a mesma periodicidade (de acordo com a Tabela 3, o número de registros por período de exibição é o dobro do mostrado).

## Resumo da metodologia estatística utilizada

Uma vez sincronizados todos os relógios dos sistemas de captura de imagens e registro de dados de cada pórtico, é possível dar início ao processo de aquisição. Com a sincronização, o processo de amostragem pode ser aplicado. Um resumo da metodologia a ser utilizada para a análise de imagens está descrito a seguir:

- 1) Para cada estrato, extrai-se um conjunto  $X$  de  $n$  registros/imagens conforme a Tabela 3;
- 2) Analisam-se todas as imagens amostradas em BackOffice e divide-se em 2 grupos: grupo conforme e não conforme de acordo com uma hierarquia pré-definida (retorno de código válido, placa legível para o humano, ilegíveis etc);
- 3) Calcula-se a média  $t = \mu_X$  para o grupo positivo (conforme);
- 4) O resultado será apresentado como a taxa medida no período:  $\mu \mp \sqrt{\frac{t(1-t)}{n}}$
- 5) Montam-se as estatísticas integrando-se dados para vários dias, horários, sítios etc. As estatísticas integradas devem levar em conta o número total ( $n$ ) de imagens analisadas com o que é possível determinar os intervalos de variação dos índices de acerto (médias).

## Representações analíticas das estimativas de notas (introdução)

Como descrito anteriormente, o BackOffice analisa todas as imagens amostradas a partir do registro e é responsável por classificar cada imagem de acordo com a estatística binomial que foi apresentada na Seção “Processo de amostragem e fundamentação estatística”. Para explicitar o processo de consolidação, utilizamos a representação:

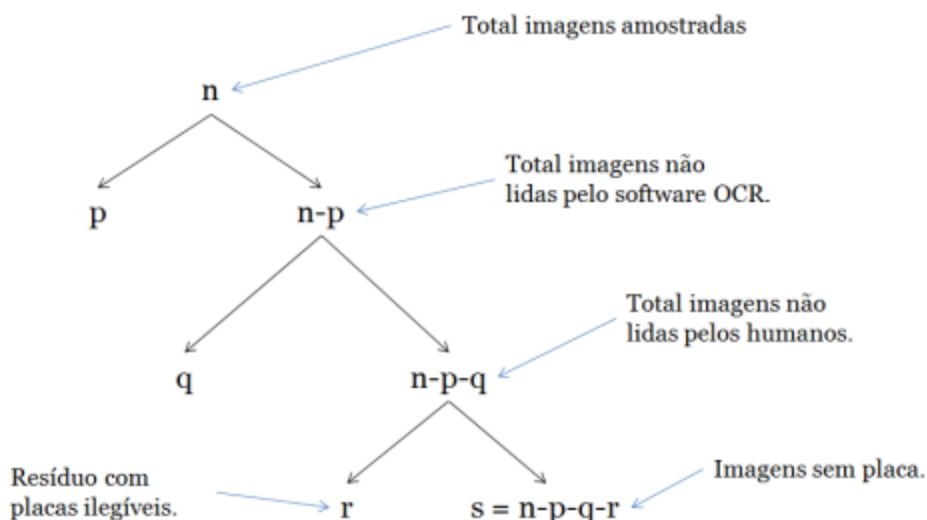


Figura 15 Hierarquia possível de problemas binomiais que resultam em relacionamento para as definições das notas usadas no processo de inspeção.

☞  $j$  = índice do sítio (pórtico, que vai de 1 a 5);

☞  $k$  = índice do período;

O processo de amostragem resultará em:

☞  $N(j,k)$  = total de registros produzidos no sítio  $j$  no período  $k$ . Assume-se que há apenas uma imagem correspondente a um registro<sup>18</sup>;

☞  $n(j,k)$  = número de registros *amostrados* no sítio  $j$  e período  $k$ . Assume-se que há pelo menos uma imagem correspondente a um registro. Aqui se observa que  $n(j,k) \leq N(j,k)$ , conforme a estratégia de amostragem utilizada.

No que segue, descrevemos, como exemplo, uma representação possível de nota, seu erro associado e sua consolidação. Por tal processo, entendemos a integração de todos os resultados após a análise em BackOffice. O processo de inspeção é representado por um diagrama de alternativas.

A Figura 15 traz um exemplo possível que já foi apresentado anteriormente<sup>19</sup>. Uma vez realizada a conferência pelo BackOffice das placas lidas automaticamente, obtém-se, para cada período no sítio  $j$ , o número total de imagens corretamente lidas pelo software OCR:

☞  $p(j,k)$  = número de imagens lidas corretamente pelo software OCR no sítio  $j$  no período  $k$ ;

Portanto, podemos definir (Figura 15):

<sup>18</sup> Durante o projeto, foi exigido dos fornecedores que apenas *uma imagem* fosse disponibilizada para cada registro. A existência de múltiplas imagens pode ser útil em BackOffice já plenamente organizado para prover redundância de análise, entretanto, no escopo e tempo de projeto, essa situação gera problemas de armazenamento e dificuldades de tratamento de múltiplas imagens na estatística final.

<sup>19</sup> Esse é apenas um exemplo para descrever o processo que foi usado para análise de dados.

$$\%ICE(j, k) = \frac{p(j, k)}{n(j, k)}$$

como sendo o percentual de imagens corretamente lidas pelo software OCR no sítio  $j$  e período  $k$ . Portanto:

$$\%INCE(j, k) = 1 - \%ICE(j, k)$$

é o percentual de imagens não lidas corretamente pelo software OCR. Nota-se que o erro de avaliação dessa nota é igual ao de  $\%ICE(j, k)$  conforme está mostrado na Figura 14.

### Cálculo do erro de uma nota e sua consolidação

Uma nota é representada como uma fração ou índice, relação entre duas quantidades dadas em termo do índice do sítio  $j$  e índice do período  $k$ ,  $\%I(j, k)$ . Representemos um índice ou nota qualquer por:

$$\%I(j, k) = \frac{\alpha(j, k)}{\beta(j, k)} \pm \sigma_I(j, k),$$

onde  $\sigma_I(j, k)$  é o desvio padrão associado e  $\beta(j, k) > \alpha(j, k)$ . Essa nota representa um índice do problema binomial e, portanto, temos:

$$\sigma_I(j, k) = \sqrt{\frac{\alpha(j, k)[\beta(j, k) - \alpha(j, k)]}{\beta(j, k)^3}},$$

O mesmo índice consolidado para o sítio,  $\%I(j)$ , será dado como a integração de todos os dados *para todos os períodos*. Portanto:

$$\%I(j) = \frac{\sum_{k=1}^{48} \alpha(j, k)}{\sum_{k=1}^{48} \beta(j, k)} \pm \sigma_I(j)$$

com

$$\sigma_I(j) = \sqrt{\frac{\sum_{k=1}^{48} \alpha(j, k)}{\sum_{k=1}^{48} \beta(j, k)^3} \left\{ \sum_{k=1}^{48} [\beta(j, k) - \alpha(j, k)] \right\}},$$

O mesmo índice, consolidado para o período,  $\%I(k)$ , será dado como a integração de todos os dados *para todos os sítios*.

$$\%I(k) = \frac{\sum_{j=1}^5 \alpha(j, k)}{\sum_{j=1}^5 \beta(j, k)} \pm \sigma_I(k),$$

com

$$\sigma_I(k) = \sqrt{\frac{\sum_{j=1}^5 \alpha(j, k)}{[\sum_{j=1}^5 \beta(j, k)]^3} \left\{ \sum_{j=1}^5 [\beta(j, k) - \alpha(j, k)] \right\}}.$$

Finalmente, um índice global, %I, representa uma nota integrada para todos os sítios e períodos:

$$\%I = \frac{\sum_{j=1}^5 \sum_{k=1}^{48} \alpha(j, k)}{\sum_{j=1}^5 \sum_{k=1}^{48} \beta(j, k)} \pm \sigma_I,$$

onde

$$\sigma_I = \sqrt{\frac{\sum_{j=1}^5 \sum_{k=1}^{48} \alpha(j, k) \left\{ \sum_{j=1}^5 \sum_{k=1}^{48} [\beta(j, k) - \alpha(j, k)] \right\}}{[\sum_{j=1}^5 \sum_{k=1}^{48} \beta(j, k)]^3}},$$

Portanto, ao invés de apresentar explicitamente as relações das frações e seus desvios-padrão associados para todas as notas, fazemos referências à necessidade de se usar as equações apresentadas nesta seção.

## Definições adicionais

Algumas definições que serão usadas posteriormente são apresentadas aqui:

$q(j, k)$  = do total de imagens que não foram lidas corretamente, é número de imagens legíveis apenas por um humano<sup>20</sup> (no BackOffice) no sítio  $j$  no período  $k$ . O valor de  $q(j, k)$  é assim o total de imagens que apenas os operadores humanos conseguiram ler.

O valor de  $q(j, k)$  é assim o total de imagens que apenas os operadores humanos conseguiram ler. Portanto (Figura 15):

$$\%ICH(j, k) = \frac{q(j, k)}{n(j, k) - p(j, k)}$$

é o índice de imagens conforme os humanos observadas no sítio  $j$  e período  $k$ . Automaticamente, o índice de imagens perdidas em BackOffice será:

$$\%INCH(j, k) = 1 - \%ICH(j, k)$$

cujos erro associado é igual ao da estimativa de  $\%ICH(j, k)$ .

---

<sup>20</sup> Como definido anteriormente, esse é o número de imagens *incorretamente* lidas pelo software OCR, que são lidas corretamente por um humano. É fácil ver que, o número total de imagens lidas corretamente por um humano no sítio  $j$  e período  $k$  é  $p(j, k) + q(j, k)$ .

Quando o operador de BackOffice se depara com uma imagem ilegível, ele pode escolher entre: imagens de placas ilegíveis ou imagem *que não contém placa*.

☞  $r(j,k)$  = total de placas ilegíveis observadas em BackOffice no sítio  $j$  para o período  $k$  do total de placas;

☞  $s(j,k)$  = total de imagens sem placa visível (nenhuma placa pode ser reconhecida na imagem<sup>21</sup>) no sítio  $j$  e período  $k$  do total de placas não lidas;

Um exemplo com simulações (geração de estatística binomial para as classes  $p$ ,  $q$ ,  $r$  e  $s$ ) pode ser visto no Apêndice 2.

### Hierarquia binomial usada e notas finais

As relações acima são apenas subsidiárias para calcular as notas finais que aqui chamamos de ‘frações’. A organização do problema conforme a Figura 15 é, como discutido anteriormente, uma das maneiras de se resolver o problema. Como explicado na seção “Discussão introdutória sobre as definições das notas de eficiência”, uma imagem pode conter um código válido. Não é difícil ver que, *se um registro contém um código válido*, então a imagem correspondente deve conter uma placa. Também, *se o registro contém um código válido*, a placa existente na imagem, muito provavelmente será lida por um humano.

A verificação desse código válido é um processo relativamente simples que pode ser automatizado. Assim, a inspeção das imagens foi realizada em um processo mais detalhado do que aquele esquematizado pela Figura 15, que não se baseia em informação a respeito do código válido. Esse novo diagrama é mostrado na Figura 16.

---

<sup>21</sup> Como dito anteriormente, isso se deve a problemas do tipo: iluminação, problema de trigger, posicionamento do veículo etc.



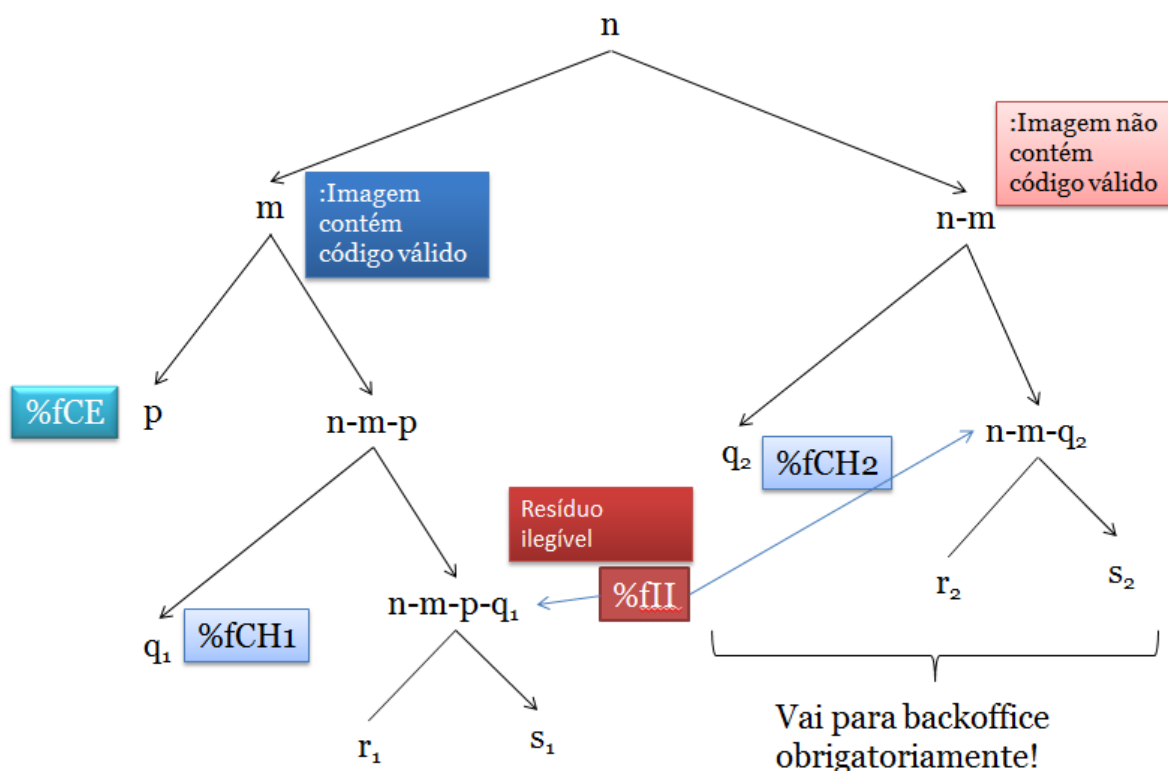


Figura 16 Hierarquia binomial usada no método de inspeção. As definições estão no texto.

Conforme a figura, definimos:

$m(j,k)$  = número de imagens para as quais um código de placa válido é retornado no sítio  $j$  e período  $k$ . Um código válido deve, necessariamente, conter uma sequência de 3 letras e 4 números.

Uma vez que há a possibilidade da imagem ter associado um código válido, dividimos em 2 classes as quantidades  $q$ ,  $r$  e  $s$ :

$q_1(j,k)$  = do total de imagens que não foram lidas corretamente, é número de imagens legíveis apenas por um humano (no BackOffice) no sítio  $j$  no período  $k$  que contém um código válido.

$q_2(j,k)$  = do total de imagens que não foram lidas corretamente, é número de imagens legíveis apenas por um humano (no BackOffice) no sítio  $j$  no período  $k$  que não contém um código válido.

$r_1(j,k)$  = total de placas ilegíveis com código válido observadas em BackOffice no sítio  $j$  para o período  $k$ ;

$r_2(j,k)$  = total de placas ilegíveis sem código válido observadas em BackOffice no sítio  $j$  para o período  $k$ ;

$s_1(j,k)$  = total de imagens sem placa (nenhuma placa pode ser reconhecida na imagem) com código válido no sítio  $j$  e período  $k$ ;

☞  $s_2(j,k)$  = total de imagens sem placa (nenhuma placa pode ser reconhecida na imagem) e sem código válido no sítio  $j$  e período  $k$ .

Não é difícil ver que

$$n = p + q_1 + q_2 + r_1 + r_2 + s_1 + s_2.$$

Com isso, podemos definir as notas como frações *em relação ao total de imagens*:

☞ %fCE = fração de imagens do total  $n$  para as quais o software OCR leu as placas nas imagens;

☞ %fCH = fração de imagens do total  $n$  para as quais apenas humanos conseguiram ler as placas nas imagens;

☞ %fPI = fração de placas do total  $n$  que são ilegíveis para os humanos (e software OCR);

☞ %fPP = fração de imagens do total  $n$  que não mostram placas (‘placas perdidas’).

☞ %fII = fração de placas perdidas do total de  $n$  imagens. Essa fração é a soma do total de placas ilegíveis e imagens sem placa e representa uma perda irreversível do sistema.

☞ %fPAL (fração de imagens ‘assertivamente’ lidas<sup>22</sup>) representa, do total de placas para as quais um código válido foi retornado, a fração de placas corretamente lidas.

☞ %fIL (fração de imagens lidas) é razão de imagens para as quais o software OCR retornou um código válido do total de imagens;

Dadas essas definições, para o sítio  $j$  no período  $k$  temos:

$$\%fCE(j, k) = \frac{p(j, k)}{n(j, k)},$$

$$\%fCH(j, k) = \frac{q_1(j, k) + q_2(j, k)}{n(j, k)},$$

$$\%fPI(j, k) = \frac{r_1(j, k) + r_2(j, k)}{n(j, k)},$$

$$\%fPP(j, k) = \frac{s_1(j, k) + s_2(j, k)}{n(j, k)},$$

$$\%fII(j, k) = \%fPI(j, k) + \%fPP(j, k),$$

---

<sup>22</sup> Isto é, a quantidade de imagens que se tem ‘convicção’ de terem sido corretamente lidas. Essa convicção é dada a partir da constatação que o software retornou uma placa válida.

$$\%fPAL(j, k) = \frac{p(j, k)}{m(j, k)},$$

$$\%fIL(j, k) = \frac{m(j, k)}{n(j, k)}.$$

Portanto:

$$\%fCE + \%fCH + \%fPI + \%fPP = 1,$$

A razão para a definição de %fPAL é que, no processo de inspeção de imagens, os relatórios de leitura contém o campo com os caracteres supostamente lidos pelo Software de OCR. É comum que fornecedores apresentem a nota de desempenho do sistema OCR em termos dessa definição (já adaptada para o sítio  $j$  e período  $k$ ). Essas definições levam a um aumento ‘artificial’ da nota de inspeção, já que  $m(j, k) \leq n(j, k)$ , ou seja,  $\%fPAL \geq \%fCE$ . Portanto, a eficiência de leitura de um sistema de OCR deve ser dada em termos de %fCE que representa a fração, do *total* de registros adquiridos, contendo imagens que foram lidas corretamente. A existência de valores  $r_1$  e  $s_1$  diferentes de zero implica que o fornecedor pode ter propositalmente chutado códigos válidos<sup>23</sup> em imagens que, ou são ilegíveis para humanos (e, portanto, para o software pelo método escolhido de análise), ou não contém placa. É importante ressaltar que os desvios padrão para todas as notas ou frações definidas acima podem ser calculados a partir das relações fornecidas na Seção “Cálculo do erro de uma nota e sua consolidação”.

Além dos índices descritos acima, também definimos variáveis para auditoria de tráfego:

☞ NT( $j, k$ ) total de *registro de tráfego* obtido por dispositivo independente (chamado de ‘aferidor de tráfego’);

☞ %IrV( $j, k$ ): índice de perda de tráfego pelo sistema de inspeção de imagens. Ele é definido por:

$$\%IrV(j, k) = \frac{NT(j, k) - N(j, k)}{NT(j, k)},$$

De acordo com essa definição temos:

- Se %IrV for *positivo*, há *perda de registro* pelo sistema;
- Se %IrV for *negativo*, há *geração de registros espúrios* no sistema (ou seja, multiplicidade de registros);

---

<sup>23</sup> Esse ‘chute’ ocorre de forma não proposital ao se permitir que outros caracteres presente na imagem sejam lidos como placas;

Embora seja possível dizer que %IrV nunca será exatamente zero para a maioria dos sistemas<sup>24</sup> e que, na verdade, haverá um valor máximo para %IrV, a presença de valores negativos é um indicador certo de falta de ajuste.

O índice de ocorrência de motos é chamado %fT e é simplesmente igual ao total observado de motos (após inspeção em BackOffice) no sítio j e período k dividido por n(j,k).

### Correção das frações pelo fluxo real

Um sistema de leitura de placas após 10 passagens iniciais apenas consegue ler corretamente 2 registros. Entretanto, o fornecedor desse sistema reporta ter recebido apenas 3 registros, com o que ele calcula a sua nota:  $2/3 = 66,6\%$ . Porém, sabemos que sua nota correta (normalizada) seria  $2/10 = 20\%$ . De forma explícita, podemos descrever a situação pela seguinte sequência de operações:

$$\frac{3}{10} + \frac{7}{10} = 1, \rightarrow \frac{2+1}{10} + \frac{7}{10} = 1, \rightarrow$$

$$\frac{\left(\frac{2}{3} + \frac{1}{3}\right)^3}{10} + \frac{7}{10} = 1, \rightarrow \left(\frac{2}{3} + \frac{1}{3}\right) \left(\frac{3}{10}\right) + \frac{7}{10} = 1$$

O fator grafado como ‘**3/10**’ é o fator de correção que deve ser multiplicado à nota (2/3) para que sua nota verdadeira seja obtida ( $66,6\% \cdot 30\% = 20\%$ ). Infelizmente ao longo do projeto, o índice de perda observado não foi homogêneo entre vários fornecedores. Por causa disso, foi necessário normalizar as frações com base nas perdas reais observadas. Isso significa que cada sítio k e período j tem um fator de correção de índice que é função de IrV(j,k) e de todas as frações inferidas:

☞ y(j,k) fator de correção por não reporte completo de registros (perdidos).

O problema da correção se deve assim ao fato de alguns fornecedores não terem reportado registros. Ou seja, os relatórios de inspeção automáticos somente contêm códigos válidos. Além disso, mesmo quando um veículo é efetivamente perdido (não há registro de imagem), um registro válido não é gerado. Para ver como foi feita a normalização consideremos a equação:

$$\%fCE + \%fCH + \%fPI + \%fPP = 1,$$

Nos casos dos fornecedores que não reportam veículos perdidos, essa equação deve ser modificada para incluir o termo %IrV  $\rightarrow$  %fPP. Se todos os registros fossem reportados, a seguinte equação seria válida para o sítio k no período j:

$$y(j, k)[\%cfCE + \%cfCH + \%cfPI] + \%IrV(j, k) = 1,$$

<sup>24</sup> Isto é, um dado sistema pode saber que perdeu um veículo e reportar isso corretamente.

onde os índices “%cf...” são as frações corrigidas. Em uma amostragem onde não se reportem veículos perdidos, as frações normalizadas são as notas efetivas do sistema e são iguais a:

$$\%cfCE(j, k) = y(jk)\%fCE(j, k),$$

$$\%cfCH(j, k) = y(jk)\%fCH(j, k),$$

$$\%cfII(j, k) = y(jk)\%fII(j, k),$$

com

$$y(j, k) = \frac{1 - IrV(j, k)}{\%fCE(j, k) + \%fCH(j, k) + \%fII(j, k)} = 1 - IrV(j, k),$$

O denominador da equação acima é a soma das notas obtidas a partir da análise dos registros que foram *efetivamente* reportados é igual a 100%. Portanto, o fator de correção é tanto menor quanto maior for o índice de perda  $IrV(j, k)$ .

Imaginemos que em um ponto há perda de 70% do fluxo. Se, com base nas imagens efetivamente reportadas temos que  $\%fCE + \%fCH + \%fII = 100\%$ , então o fator de normalização será:

$$y = \frac{1 - 0,7}{1} = 30\%$$

Portanto, se  $\%fCE = 75\%$ , por exemplo,

$$\%cfCE(j, k) = 30,0\% * 75\% = 22,5\%,$$

que é a nota real de desempenho do sistema.

### **Exemplo de cálculo de frações de classes de leitura (GRU 23/8/2012)**

O exemplo a seguir serve para apresentar um caso real para a estatística apresentada anteriormente. Os números resultantes foram obtidos de uma amostra muito pequena inicial *quando o fornecedor reconhecidamente ainda estava realizando ajustes no sistema* e não representam assim as notas finais. Como o número de imagens por período é pequeno, reunimos dois grandes períodos conforme descrito a seguir.

#### Período diurno (das 12:00 as 18:00 e das 07:00 as 11:00)

Descrição dos resultados: o total de imagens analisadas como teste da metodologia foi de 117 imagens, onde se excluíram motos (Figura 17). *Do total de imagens*, aproximadamente 81,2% (%fIL) teve retorno do algoritmo OCR com algum código de placa válido. *Do total de imagens contendo código válido*, aproximadamente 90,5%(fPAL) foram lidos corretamente. Entretanto, *do total de imagens*, apenas

73,5% (%fCE) foram lidas corretamente, 18,8% (%fICH) foram lidas apenas por humanos, 4,3%(%fPI) contém placas ilegíveis e 3,4% (%fPP) não contém placa alguma. Isso significa que, *do total de imagens*, 7,7% (%fII) foram perdidas irremediavelmente.

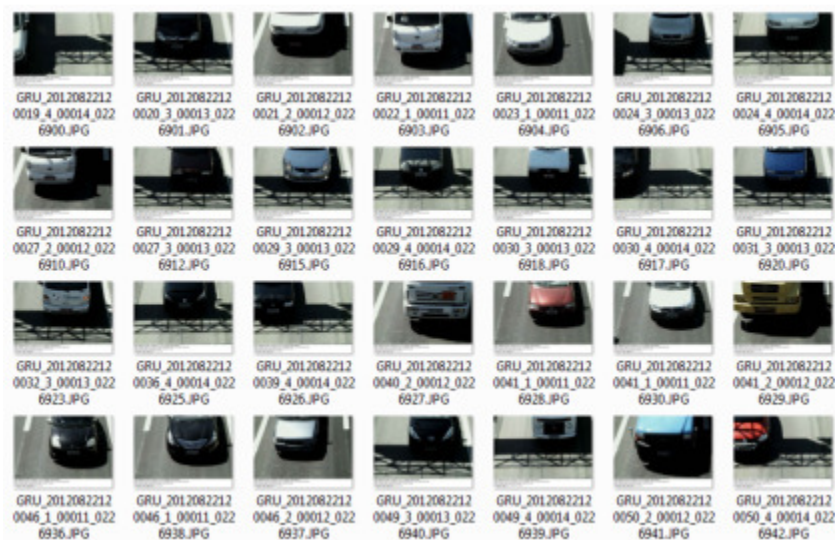


Figura 17 Imagens diurnas frontais adquiridas em GRU em 23/8/2012.

1. Total de imagens no período (n): 117 (excluído motos)
2. Total de imagens lidas corretamente pelo software OCR (p): 86
3. Total de imagens com placa que só podem ser lidas por humanos (q): 22
4. Total de imagens com placas ilegíveis (r): 5
5. Total de imagens que não contém placa (s): 4
6. Total de imagens assertivamente lidas (m): 95
7. Total de imagens ilegíveis (soma de 4 e 5): 9

### Frações

- %fCE =  $73,5 \pm 4,1\%$ pp (fração de imagens com placa lidas corretamente pelo software de OCR do total)
- %fCH =  $18,8 \pm 3,6\%$ pp (fração de imagens com placas que somente humanos podem ler do total)
- %fPI =  $4,3 \pm 1,9\%$ pp (fração de imagens com placas ilegíveis do total)
- %fPP =  $3,4 \pm 1,7\%$ pp (fração de imagens sem placa do total)

### Outros índices

- %fPAL =  $90,5 \pm 3,0\%$ pp (fração de imagens com placa assertivamente lidas)
- %fII =  $7,7 \pm 2,5\%$ pp (fração total de imagens ilegíveis ou perdidas)
- %fIL =  $81,2 \pm 3,6\%$  (fração de imagens lidas)

### Período noturno (18:00 as 06:00)

Descrição dos resultados: o total de imagens analisados como teste da metodologia foi de 399 imagens, onde se excluíram motos (Figura 18). *Do total de imagens*,



aproximadamente 42,9% (%fIL) teve retorno do algoritmo OCR com algum código de placa válido. *Do total de imagens contendo código válido*, aproximadamente 91,8%(fPAL) foram lidas corretamente. Entretanto, *do total de imagens*, apenas 39,3% (%fCE) foram lidas corretamente, 19,5% (%fICH) foram lidas apenas por humanos, 40,4%(%fPI) contém placas ilegíveis e 0,8% (%fPP) não contém placa alguma. Isso significa que, *do total de imagens*, 41,1% (%fII) são placas perdidas.

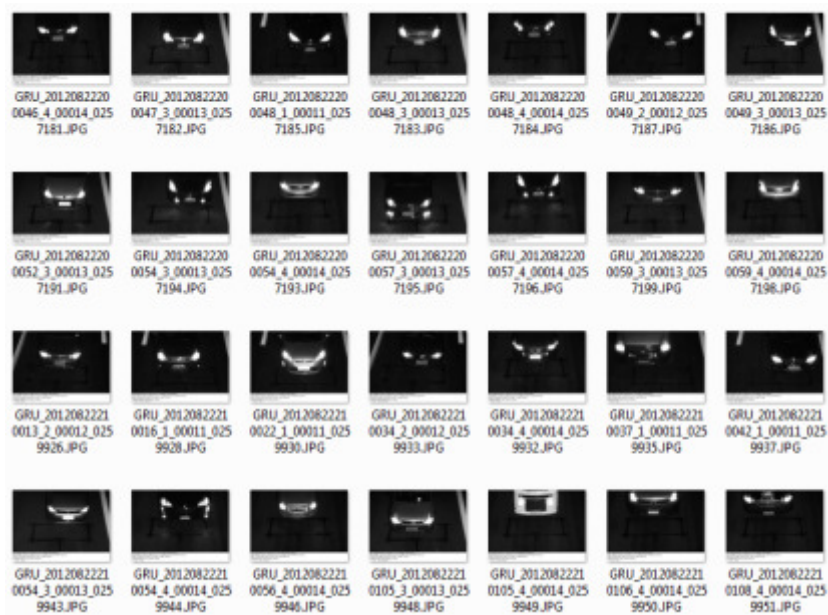


Figura 18 Imagens noturnas frontais adquiridas em GRU em 23/8/2012

1. Total de imagens no período (n): 399 (excluído motos)
2. Total de imagens lidas corretamente pelo software OCR (p): 157
3. Total de imagens com placa que só podem ser lidas por humanos (q): 78
4. Total de imagens com placas ilegíveis (r): 161
5. Total de imagens que não contém placa (s): 3
6. Total de imagens assertivamente lidas (m): 171
7. Total de imagens ilegíveis (soma de 4 e 5): 164

### Frações

- %fCE =  $39,3 \pm 2,4\%$ pp (fração de imagens com placa lidas corretamente pelo software de OCR do total)
- %fCH =  $19,5 \pm 2,0\%$ pp (fração de imagens com placas que somente humanos podem ler do total)
- %fPI =  $40,4 \pm 2,5\%$ pp (fração de imagens com placas ilegíveis do total)
- %fPP =  $0,8 \pm 0,4\%$ pp (fração de imagens sem placa do total)

### Outros índices

- %fPAL =  $91,8 \pm 2,1\%$ pp (fração de imagens com placa assertivamente lidas)
- %fII =  $41,1 \pm 2,5\%$ pp (fração total de imagens ilegíveis ou perdidas)
- %fIL =  $42,9 \pm 2,5\%$  (fração de imagens lidas)



Observe que, em ambos os períodos (diurno e noturno), as notas %fPAL estão acima de 90%. Entretanto, o caso noturno considerado apresenta um baixo índice de aproveitamento do sistema OCR (apenas 42,9% das imagens retornaram código válido), o que implicou em alto índice de imagens com placa ilegíveis (%fPI). Isso se deveu a um problema de iluminação das placas no período diurno onde o sistema estava ajustado apenas para captura ótima de placas sem camada reflexiva. Depois disso, um ajuste foi feito o que possibilitou ler de forma otimizada tanto placas reflexivas como não reflexivas. Esse exemplo demonstra como o conjunto de frações utilizadas consegue caracterizar bem o desempenho do sistema de forma específica e não apenas global.

### Processo de inspeção de imagens em BackOffice

A parte anterior apresentou a teoria que fornece as referências analíticas para se instaurar um processo de inspeção. Essa, por sua vez, é feita por operadores humanos idealmente isentos e treinados que analisam imagem a imagem com o objetivo de se determinar – pelo acúmulo de análises sucessivas – os índices e as frações definidas anteriormente.

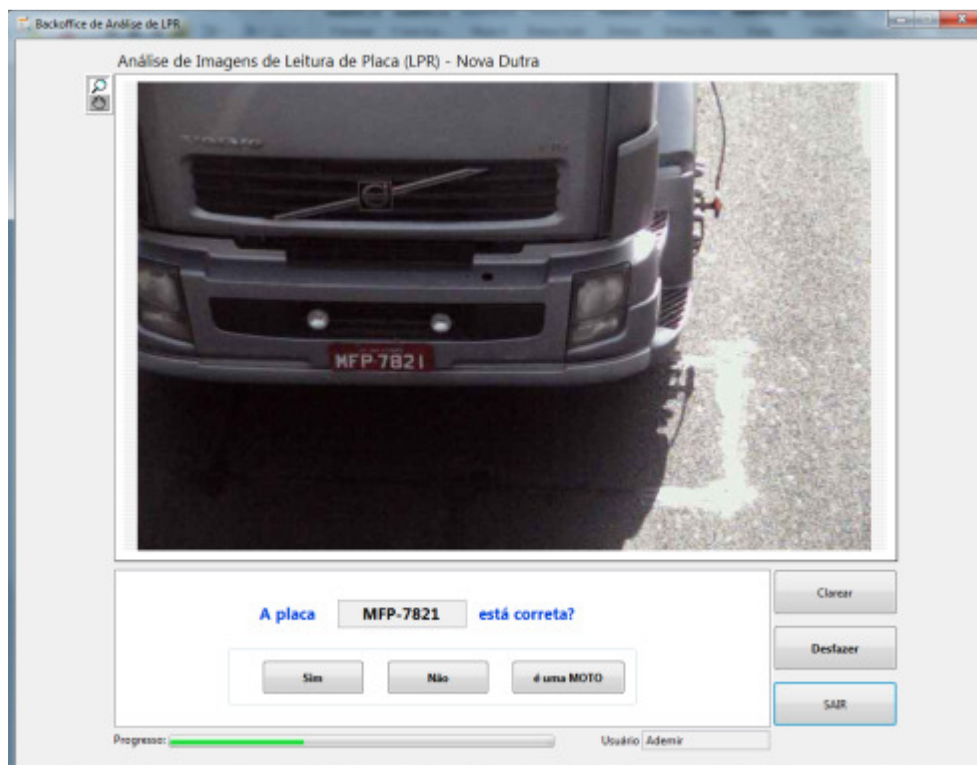


Figura 19 Interface usada para inspeção de imagens em BackOffice. Essa tela é a primeira no caso de um veículo para o qual o sistema retornou um código válido.

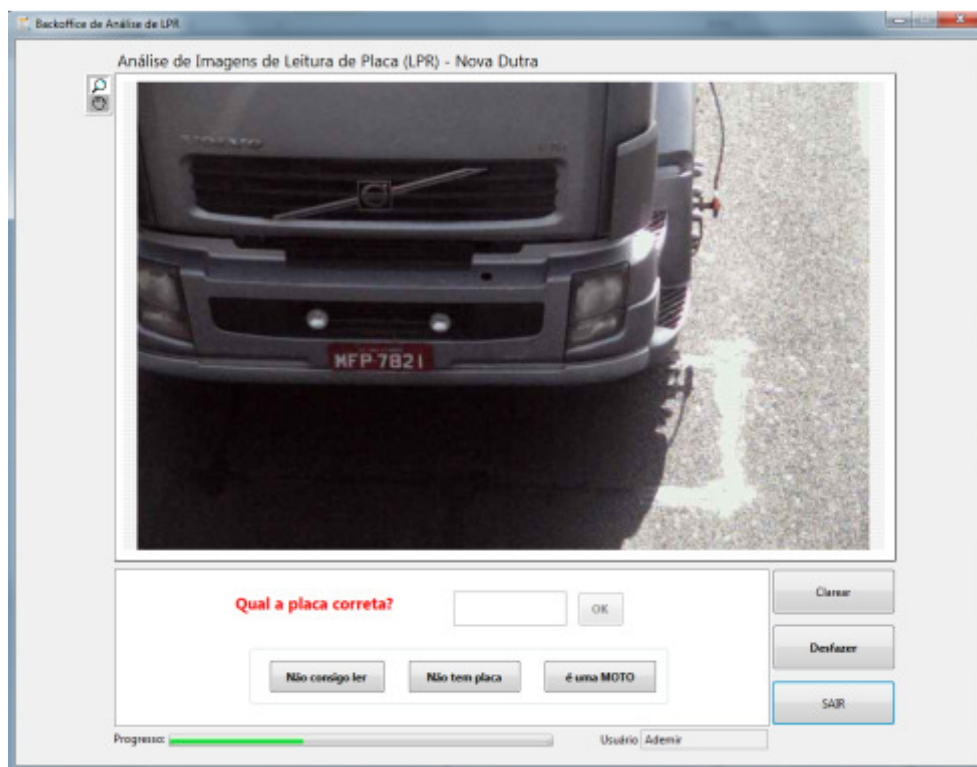


Figura 20 Interface usada para inspeção de imagens em BackOffice. Essa é a segunda tela no caso de um veículo para o qual o sistema retornou um código válido.

A Figura 19 e Figura 20 ilustram a interface de inspeção de imagens que foi usada para o BackOffice. Essa interface permite:

- Verificar se a imagem mostrada contém uma placa:
  - Se ela contém uma placa, verificar se o código escrito no campo corresponde exatamente ao código da placa na imagem;
    - Se o código coincide, o botão ‘SIM’ é pressionado;
    - Se não coincide, o botão ‘NÃO’ é pressionado e o código lido pelo humano deve ser preenchido em um campo específico (Figura 20);
  - Se ela contém uma placa que não pode ser lida na imagem (placa ilegível), o botão ‘NÃO CONSIGO LER’ é pressionado (placa visível, mas ilegível, Figura 20);
- Se ela não contém uma placa, o botão ‘NÃO TEM PLACA’ é pressionado;
- Caso a ocorrência seja de moto, o botão ‘É UMA MOTO’ deve ser pressionado.

Internamente, a estatística é gerada segregando-se  $n(j,k)$  imagens<sup>25</sup> em classes correspondentes às letras p, q, r e s, cujo relacionamento com a estatística binomial é mostrada na Figura 16. O processo gera assim as matrizes  $p(j,k)$ ,  $q(j,k)$ ,  $r(j,k)$  e  $s(j,k)$  para o conjunto  $n(j,k)$  de imagens, que endereça as proporções discutidas

<sup>25</sup> Como j máximo é 5 e k máximo é 48, há 240 grupos diferentes de imagens.

anteriormente a partir das quais se podem gerar todos os índices e frações e seus desvios padrão associados.

O processo de inspeção em BackOffice foi concebido de forma a evitar ‘biasing’, isto é, a formação de tendências de atribuição característica de cada inspetor em BackOffice entre inspetores. Para isso, um procedimento de embaralhamento de dados (*DATA Shuffling*) foi concebido no processo de consolidação de dados (Figura 21) de forma que um arquivo único fosse gerado com referência a imagens fornecidas. Esse arquivo único contém entradas de diferentes sítios e os inspetores analisam imagens de todos os fornecedores *sem qualquer identificação de origem* (ou seja, os inspetores não sabem a origem ou sítio onde a imagem foi adquirida). Dessa forma, evita-se a formação de ‘bias’ (por auto treinamento ou especialização do modo de inspeção a uma determinada fonte de dados) e reduz-se o stress dos inspetores por analisar dados provenientes de pontos problemáticos<sup>26</sup>. Depois da finalização do processo de inspeção, os dados são separados (via outro procedimento de *DATA Unshuffling*) e as estatísticas específicas por sítio de aquisição são geradas.

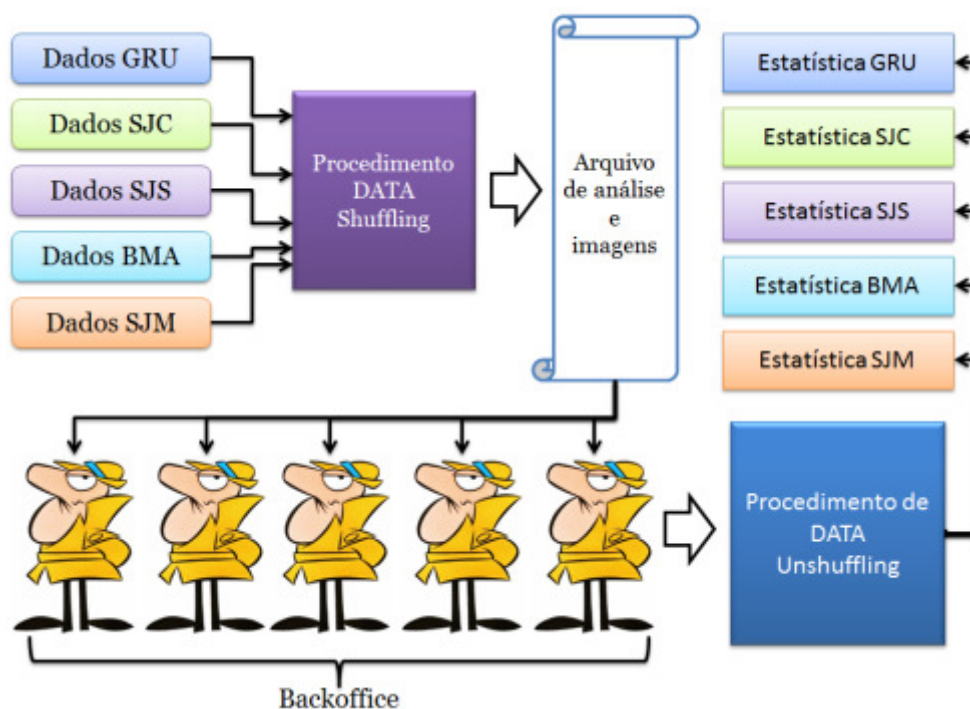


Figura 21 Diagrama descritivo do processo de análise em BackOffice. Os arquivos de dados não são enviados aos inspetores de forma separada e, muito menos com identificação de origem. Antes disso, um procedimento de ‘embaralhamento’ foi adotado que é seguido por um ‘desembaralhamento’ antes da determinação das estatísticas específicas.

Além do levantamento das estatísticas das classes, para dimensionamento de esforço em RH de BackOffice, é feito o cálculo do tempo médio de inspeção por placa por pessoa. Isso é conseguido medindo-se o tempo com que cada indivíduo inspeciona  $n(j,k)$  imagens e tirando-se a média e o desvio padrão para as  $n(j,k)$  imagens. Como o total de imagens está dividido em vários sítios e períodos, é possível

<sup>26</sup> Por exemplo, pórtilos cujos sistemas exigem esforço maior de BackOffice para análise.

com isso verificar eventuais variações no tempo de inspeção (entre, por exemplo, imagens diurnas e noturnas) e assim por diante.

Da Figura 16 é possível gerar uma sequência de perguntas que compõem um ‘wizard’<sup>27</sup> da interface de BackOffice. Essa sequência está representada pelo conjunto de instruções conforme abaixo:

```

→Abre o i-ésimo registro do banco;
→SE (imagem contém um registro válido) ENTÃO:
    →Exibe imagem;
    →Pergunta: O código está conforme a placa na imagem?
    SE (sim) ENTÃO:
        →Atualiza contador p
    CASO CONTRÁRIO:
        → Pergunta: Você consegue ler a placa?
        SE (sim) ENTÃO:
            →Digita código da placa
            →Atualiza contador q1
        CASO CONTRÁRIO:
            →Pergunta: Você consegue ver uma placa na
            imagem?
            SE (sim) ENTÃO:
                →Atualiza contador r1
            CASO CONTRÁRIO:
                →Atualiza contador s1
→CASO CONTRÁRIO:
    → Pergunta: Você consegue ler a placa?
    SE (sim) ENTÃO:
        →Digita código da placa
        →Atualiza contador q2
    CASO CONTRÁRIO:
        →Pergunta: Você consegue ver uma placa na imagem?
        SE (sim) ENTÃO:
            →Atualiza contador r2
        CASO CONTRÁRIO:
            →Pergunta: A imagem contém uma moto?
            SE (sim) ENTÃO
                →Atualiza contador de motos
            CASO CONTRÁRIO:
                →Atualiza contador s2
→ Atualiza i-ésimo registro
  
```

Como apresentado anteriormente, ocorreu também a necessidade de segregar imagens contendo motos (desde que visíveis). Registros contendo motos não devem (idealmente) gerar códigos válidos (exceto pela aquisição de imagem traseira, o que foi segregado de forma específica, ver Seção “Leitura automática de motos”). Portanto, um contador especial foi criado para segregar imagens contendo motos.

<sup>27</sup> Um ‘wizard’ é um algoritmo que facilita a inserção de dados em uma interface.

## Regras adicionais para inspeção das imagens



Figura 22 Imagem da traseira de um caminhão contendo um adesivo com informação da placa. Não há, entretanto, nenhuma placa visível.

O processo de análise só está completo se regras adicionais são especificadas para possibilitar que o inspetor corretamente classifique as imagens. Essas regras são:

1. A leitura automática somente será considerada correta (e, portanto, pertencente à classe  $p(j,k)$ ) quando todos os caracteres fornecidos pela inspeção forem coincidentes sequencialmente com os observados na placa contida na imagem;
2. A imagem fornecida deve conter uma placa legível para que seja corretamente classificada. Embora imagens sem placa legível estejam previstas na estatística, é possível que o responsável por cada sítio – ao preparar as imagens – utilize mais de uma imagem para extrair (corretamente) o código da placa. Cria-se uma situação em que um código é corretamente lido, mas a imagem enviada para inspeção não pode ser lida por um humano (como, por exemplo, na presença de neblina onde um processo de amostragem de imagens pode resultar na leitura correta). Todo o processo de inspeção em BackOffice é baseado em inspetores humanos que são admitidos como superiores ao sistema automático. Reitera-se que

*todo processo de inspeção de BackOffice depende exclusivamente da existência de placas legíveis nas imagens*<sup>28</sup>;

3. Não será permitido o encaminhamento de mais de uma imagem por placa. Por causa de limitações de memória (dada o fluxo esperado nos 5 sítios de aquisição) e necessidade de se fornecer uma estatística confiável por meio de amostragem aleatória entre sítios, múltiplas imagens e registros não serão considerados no processo de inspeção<sup>29</sup>;
4. Como parte dos dados envolve aquisição de imagem de traseira de veículos, a existência de caracteres contendo a informação da placa em outras partes do veículo que não sejam placas, por exemplo, adesivos traseiros (DENATRAN, 2010) *não serão considerados como leituras legítimas* e, portanto, contabilizados na estatística  $s(j,k)$ . Um exemplo pode ser visto na Figura 22. É importante considerar que a regra que permitiu o uso de adesivos para visualização de placas foi indeferida (DENATRAN, 2011).

## **Análise de fluxo, velocidade e densidade de veículos**

Um sítio durante o projeto onde foi possível medir a densidade de veículos da forma precisa (o que envolve determinação da velocidade instantânea dos veículos durante a passagem) foi o pórtico de Guarulhos. Para isso, foi instalado, além do contador de tráfego<sup>30</sup> um conjunto de laços adicionais para medida da velocidade. Dessa forma, no ponto de Guarulhos com cerca de 8 pistas de aferição, foi instalado um conjunto de 16 laços para medida de velocidade, além da contagem de veículos e função de disparo das câmeras.

Definimos os seguintes termos:

- ☞ Fluxo de veículos ( $f$ ): é o volume de veículos (por hora, p. ex.) retornado pela integração de dados de contagem em um intervalo de tempo pré-definido;
- ☞ Velocidade de veículos ( $v$ ): é a *velocidade instantânea* dos veículos sob o pórtico;
- ☞ Densidade de veículos ( $\eta$ ): é o número médio de veículos ocupando a via por unidade de comprimento. Quanto maior a densidade, maior é o grau de congestionamento.

Não é difícil ver que (Gazis, 2002), considerando médias de valores para todas as quantidades abaixo:

$$f = v\eta,$$

---

<sup>28</sup> Portanto, eventuais ocorrências de campos lidos corretamente sem a correspondente placa legível serão encaminhadas à estatística do botão correspondente à classe  $s(j,k)$ .

<sup>29</sup> É fácil ver que, embora seja possível segregar imagens múltiplas em sequência, com a amostragem aleatória, imagens e registros múltiplos podem ser encaminhados de forma separada para inspeção em BackOffice, o que causa aumento ou redução artificial nas notas de desempenho do sistema para fornecedores que enviem registros múltiplos.

<sup>30</sup> Um contador de laço simples apenas contabiliza o total de veículos dentro de um intervalo pré-definido.



Portanto, é possível conhecer a densidade (em veículos por quilômetro) quando se dispõe da informação da velocidade.

É importante considerar que a análise de densidade é relevante para a determinação da dependência das notas com o estado do tráfego (ou seja, se há ou não ocultamento devido a veículos altos, depende de densidades elevadas). Há medidas mais refinadas que poderiam ser feitas na determinação de notas de eficiência como função da velocidade instantânea do veículo, além do volume do tráfego. O sistema de laços duplos também permitiu obter uma estimativa da classificação do veículo por comprimento, já que essa dimensão pode ser calculada uma vez que a velocidade instantânea é conhecida. A aferição de comprimento é importante para se saber a distância média entre veículos que depende, como é fácil ver, não só do número de veículos como também seu comprimento.

Uma análise detalhada desse fluxo, densidade e velocidades instantâneas para o sítio GRU pode ser encontrado no Apêndice 3.

## **Descrição de resultados (análise de dados e tabelas com frações)**

No que segue, apresentamos resultados preliminares de determinação de frações de cada sistema. Esses relatos formam uma *descrição histórica* do andamento ou progresso do projeto.

### **Análise de dados (5/9/2012)**

Período analisado: 4 últimos dias disponíveis no servidor localizado na sede da Nova Dutra em St. Isabel/SP, capturados antes do dia 04/09/2012.

Problemas gerais encontrados:

- Períodos e, em alguns casos, dias inteiros faltando no servidor;
- Latência<sup>31</sup> entre o ponto de coleta e o servidor esteve elevada em alguns trechos, em até 7 dias (SJS e SJM). O desejado seria uma hora no máximo, depois do arquivo ser disponibilizado no ponto de coleta.
- Inconsistência de informações: arquivos incompletos, como foto sem dados;
- Registros duplicados: mais de um registro e foto para um mesmo veículo (em SJC e SJM);
- Dados não confiáveis em alguns trechos e períodos (em BMA não foram encontrados veículos não identificados ou sem placa);
- Até a data analisada indicada o fluxo local por período e pista para medição e validação dos períodos não foi fornecido;

---

<sup>31</sup> Diz respeito entre a data quando uma determinada aquisição foi feita no ponto e a data quando os dados foram disponibilizados no servidor para análise. Uma latência grande permite que os dados sejam manipulados antes de serem disponibilizados.



- Ausência de arquivos de LOG dificultou o levantamento de intervenções nos equipamentos (SJS);
- Presença de arquivos de LOG com nomenclatura incorreta (em SJC);
- Foi observado que equipamentos que capturam as placas traseiras estavam configurados/otimizados para a leitura do adesivo ao invés da placa.
- Fotos gravadas com tamanho muito grande (em SJS), além do necessário, chegando a 250KBytes em alguns casos.

Arquivos de LOG são arquivos texto que descrevem o status do equipamento e que permite saber se algum problema ocorreu durante o processo de aquisição. Alguns fornecedores não tinham disponibilizado esses arquivos até a data 5/09.

É bastante provável que os problemas de fragmentação de arquivos (tanto de Log como dos registros) tenham ocorrido em virtude do processo de ‘filetransfer’ implementado<sup>32</sup>, que não foi capaz de lidar com o fluxo muito grande de informação gerado pela operação contínua de todos os pontos.

### Frações descritas no relatório preliminar de 10/09/2012

Mesmo com as dificuldades encontradas e relatadas em 5/9, um relatório de cálculo preliminar de índices foi feito com base em 200 imagens por sítio e período (dividido simplesmente em ‘manhã’, ‘tarde’ e ‘noite’). Esse relatório seguiu exigência de cronograma que foi acertado previamente, com dada para disponibilização em 10/09. Por causa das dificuldades descritas anteriormente, apenas foram disponibilizadas notas para os pontos GRU, SJS (parcialmente, apenas o ponto que permanece na parte Sul foi determinado) e SJM conforme descritas abaixo.

GRU 01/09/2012								
MANHÃ			TARDE			NOITE		
Frações		erros	Frações		erros	Frações		erros
%fCE	78,00%	2,93%	%fCE	74,50%	3,08%	%fCE	73,87%	3,11%
%fCH	3,50%	1,30%	%fCH	4,00%	1,39%	%fCH	9,55%	2,08%
%fPI	8,50%	1,97%	%fPI	16,00%	2,59%	%fPI	11,06%	2,22%
%fPP	2,00%	0,99%	%fPP	3,00%	1,21%	%fPP	2,01%	0,99%
%fPAL	94,55%	1,77%	%fPAL	97,39%	1,24%	%fPAL	95,45%	1,62%
%fII	10,50%	2,17%	%fII	19,00%	2,77%	%fII	13,07%	2,38%
%fIL	82,50%	2,69%	%fIL	76,50%	3,00%	%fIL	77,39%	2,96%
%Motos	8,00%	1,92%	%Motos	3,27%	1,26%	%Motos	4,55%	1,48%

Figura 23 Frações e desvios padrão associados para GRU em 1/9/2012.

<sup>32</sup> Esse é o sistema responsável por operar a transferência de informação entre cada ponto e o servidor central na sede da Nova Dutra em Sta. Isabel.

SJS 26/08/2012								
MANHÃ			TARDE			NOITE		
Frações		erros	Frações		erros	Frações		erros
%fCE	73,50%	3,12%	%fCE	82,00%	2,72%	%fCE	47,00%	3,53%
%fCH	9,50%	2,07%	%fCH	5,50%	1,61%	%fCH	5,50%	1,61%
%fPI	7,00%	1,80%	%fPI	3,50%	1,30%	%fPI	11,00%	2,21%
%fPP	2,00%	0,99%	%fPP	7,00%	1,80%	%fPP	35,00%	3,37%
%fPAL	86,47%	2,66%	%fPAL	91,11%	2,22%	%fPAL	54,34%	3,88%
%fII	9,00%	2,02%	%fII	10,50%	2,17%	%fII	46,00%	3,52%
%fIL	85,00%	2,52%	%fIL	90,00%	2,12%	%fIL	86,50%	2,42%
%Motos	2,94%	1,19%	%Motos	2,22%	1,04%	%Motos	1,16%	0,76%

Figura 24 Frações e desvios padrão associados para SJS em 26/8/2012.

SJM 25/08/2012								
MANHÃ			TARDE			NOITE		
Frações		erros	Frações		erros	Frações		erros
%fCE	6,00%	1,68%	%fCE	7,00%	1,80%	%fCE	1,00%	0,70%
%fCH	30,00%	3,24%	%fCH	29,00%	3,21%	%fCH	7,00%	1,80%
%fPI	22,50%	2,95%	%fPI	33,00%	3,32%	%fPI	11,00%	2,21%
%fPP	41,50%	3,48%	%fPP	31,00%	3,27%	%fPP	81,00%	2,77%
%fPAL	80,00%	3,11%	%fPAL	77,78%	3,24%	%fPAL	25,00%	3,37%
%fII	64,00%	3,39%	%fII	64,00%	3,39%	%fII	92,00%	1,92%
%fIL	7,50%	1,86%	%fIL	9,00%	2,02%	%fIL	4,00%	1,39%
%Motos	0,00%	0,00%	%Motos	0,00%	0,00%	%Motos	0,00%	0,00%

Figura 25 Frações e desvios padrão associados para GRU em 25/8/2012

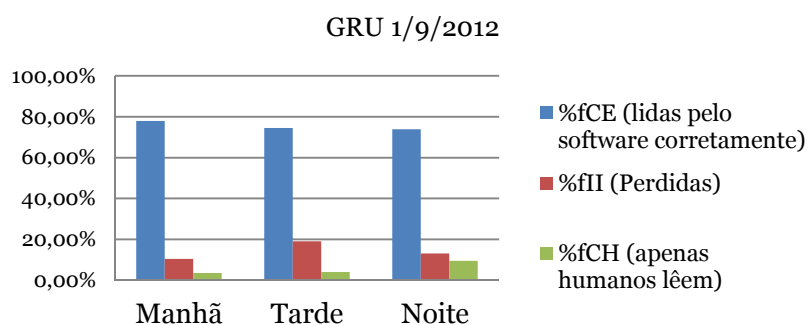


Figura 26 Frações do sítio GRU em 1/9/2012.

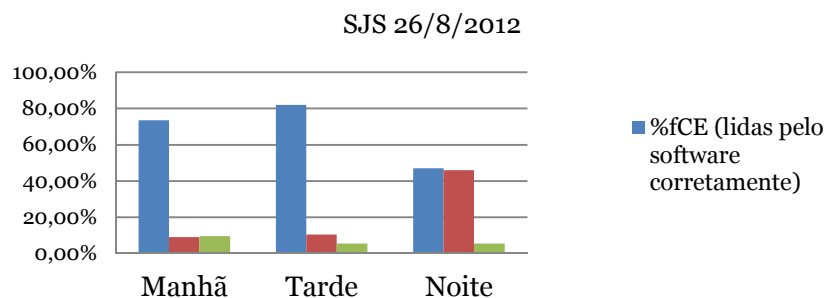


Figura 27 Frações do sítio SJS em 26/8/2012.

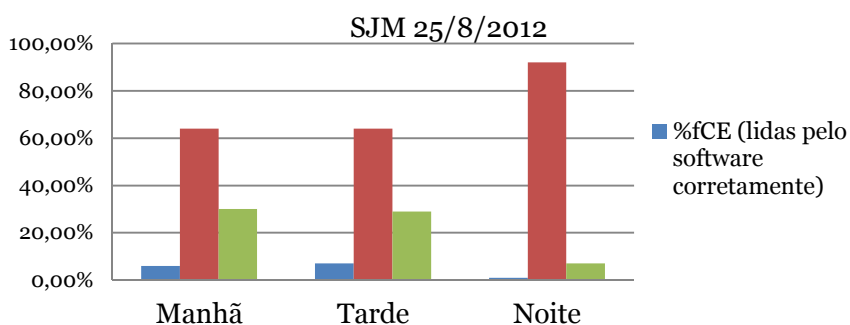


Figura 28 Frações do sítio SJM em 25/8/2012.

### Análise de dados das 14:00 as 21:00 em 6/9, 7/9 e 8/9/2012.

Adicionalmente, um conjunto limitado de aquisições foi feito nos dias 6/9, 7/9 e 8/9 para os quais se selecionou aqui a exibição de *frações integradas* para um conjunto restrito de apenas 2 sítios: GRU e SJS, ou seja, para os dias 6, 7 e 8 foram somadas, período a período, das 14:00 as 21:00 os registros e análises, de forma que fosse composta a primeira descrição temporal de variação dos índices. Não houve comparação com os níveis de tráfego por conta da limitação de pontos de aquisição e períodos. Os resultados integralizados podem ser vistos na Figura 29 e Figura 30.

GRU/HR	%fCE	%fCH	%fPI	%fPP	%fM	SfCE	SfCH	SfPI	SfPP	SfM	%fPAL	%fIL	%fII	SfPal	SfIL
14	73,2%	3,7%	6,1%	13,4%	3,7%	4,9%	2,1%	2,6%	3,8%	2,1%	95,2%	76,8%	19,5%	2,7%	4,7%
15	67,0%	4,5%	12,5%	9,1%	6,8%	5,0%	2,2%	3,5%	3,1%	2,7%	93,7%	71,6%	21,6%	3,1%	4,8%
16	74,1%	7,4%	6,2%	6,2%	6,2%	4,9%	2,9%	2,7%	2,7%	2,7%	90,9%	81,5%	12,3%	3,5%	4,3%
17	66,3%	5,8%	15,1%	9,3%	3,5%	5,1%	2,5%	3,9%	3,1%	2,0%	87,7%	75,6%	24,4%	4,1%	4,6%
18	60,7%	7,1%	11,9%	9,5%	10,7%	5,3%	2,8%	3,5%	3,2%	3,4%	98,1%	61,9%	21,4%	1,9%	5,3%
19	65,4%	6,4%	21,8%	3,8%	2,6%	5,4%	2,8%	4,7%	2,2%	1,8%	86,4%	75,6%	25,6%	4,5%	4,9%
20	61,0%	13,0%	15,6%	6,5%	3,9%	5,6%	3,8%	4,1%	2,8%	2,2%	85,5%	71,4%	22,1%	4,8%	5,1%
21	67,1%	9,2%	9,2%	13,2%	1,3%	5,4%	3,3%	3,3%	3,9%	1,3%	91,1%	73,7%	22,4%	3,8%	5,1%

Figura 29 Frações e erros associados para o sítio GRU em 6/9, 7/9 e 8/9 integrados como função do período das 14:00 as 21:00.

SJS/HR	%fCE	%fCH	%fPI	%fPP	%fM	SfCE	SfCH	SfPI	SfPP	SfM	%fPAL	%fIL	%fII	sfPal	sfIL
14	89,0%	2,4%	2,4%	3,7%	2,4%	3,5%	1,7%	1,7%	2,1%	1,7%	94,8%	93,9%	6,1%	2,5%	2,6%
15	88,7%	2,8%	8,5%	0,0%	0,0%	3,8%	2,0%	3,3%	0,0%	0,0%	91,3%	97,2%	8,5%	3,4%	2,0%
16	88,2%	5,3%	1,3%	0,0%	5,3%	3,7%	2,6%	1,3%	0,0%	2,6%	94,4%	93,4%	1,3%	2,7%	2,8%
17	84,0%	2,7%	9,3%	2,7%	1,3%	4,2%	1,9%	3,4%	1,9%	1,3%	86,3%	97,3%	12,0%	4,0%	1,9%
18	71,6%	11,1%	13,6%	1,2%	2,5%	5,0%	3,5%	3,8%	1,2%	1,7%	87,9%	81,5%	14,8%	4,0%	4,3%
19	67,1%	10,1%	15,2%	1,3%	6,3%	5,3%	3,4%	4,0%	1,3%	2,7%	91,4%	73,4%	16,5%	3,7%	5,0%
20	53,8%	15,0%	21,3%	6,3%	3,8%	5,6%	4,0%	4,6%	2,7%	2,1%	76,8%	70,0%	27,5%	5,6%	5,1%
21	72,7%	9,1%	9,1%	4,5%	4,5%	4,7%	3,1%	3,1%	2,2%	2,2%	90,1%	80,7%	13,6%	3,5%	4,2%

Figura 30 Frações e erros associados para o sítio SJS em 6/9, 7/9 e 8/9 integrados como função do período das 14:00 as 21:00.

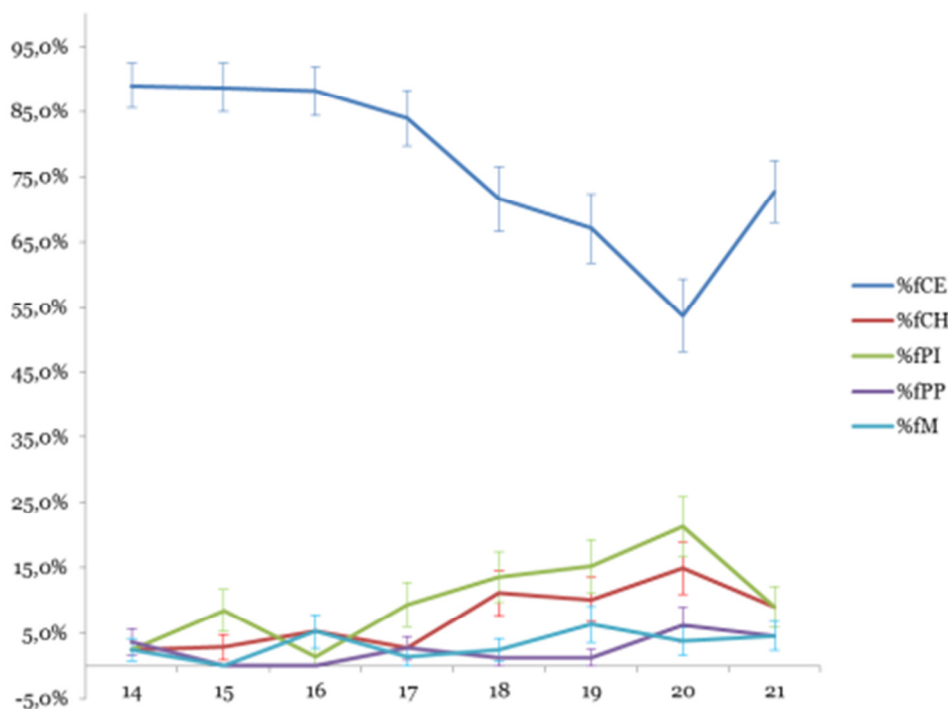


Figura 31 Frações e erros associados para o sítio GRU em 6/9, 7/9 e 8/9 integrados como função do período das 14:00 as 21:00.

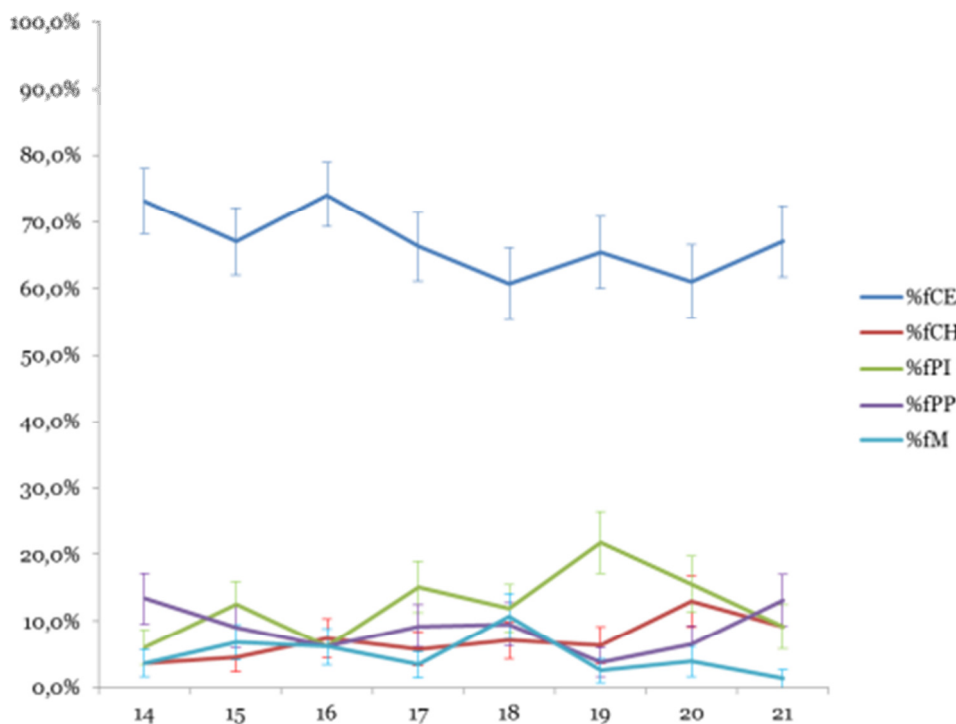


Figura 32 Frações e erros associados para o sítio SJS em 6/9, 7/9 e 8/9 integrados como função do período das 14:00 às 21:00.

A primeira parte das frações pode ser vista em forma gráfica (evolução temporal) na Figura 31 e Figura 32. As frações %fPAL e %fIL (índices de acerto de registros válidos como função do número total de imagens válidas geradas) podem ser vistos na Figura 33 e Figura 34.

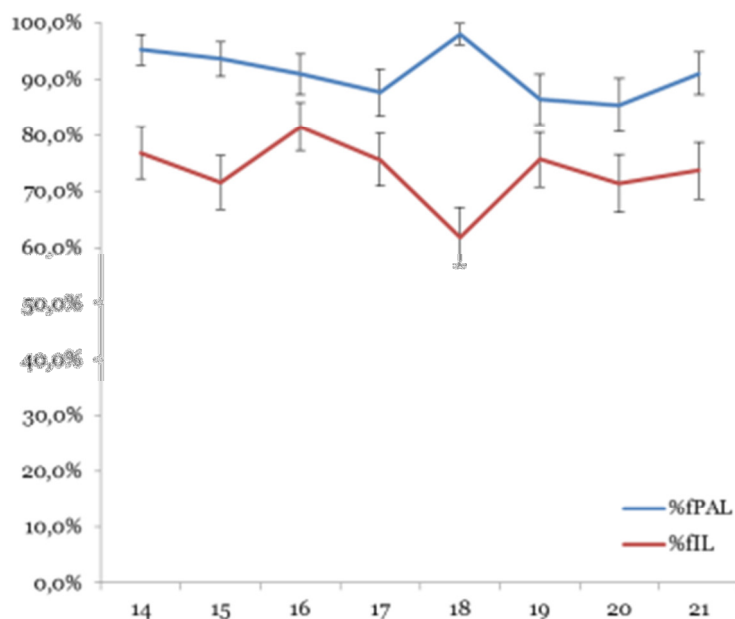


Figura 33 Frações e erros associados para o sítio GRU de %fPAL e %fIL em 6/9, 7/9 e 8/9 integrados como função do período das 14:00 às 21:00.

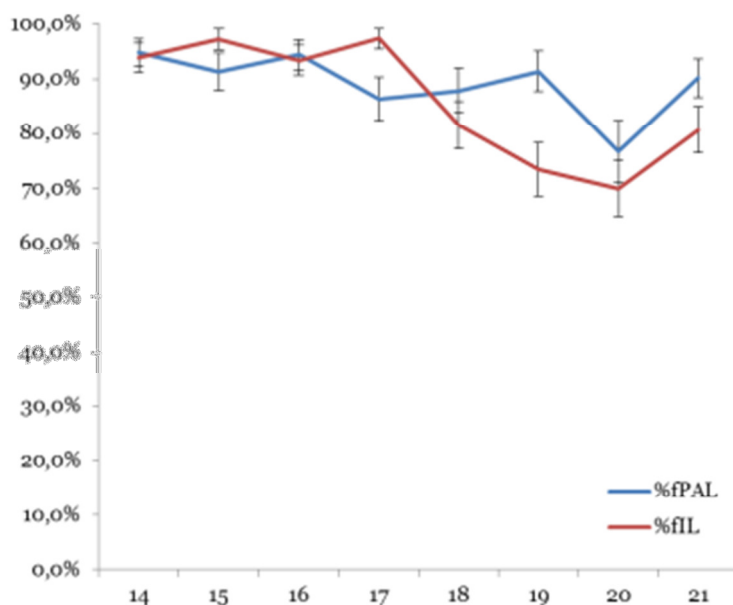


Figura 34 Frações e erros associados para o sítio SJS de %fPAL e %fIL em 6/9, 7/9 e 8/9 integrados como função do período das 14:00 as 21:00.

### **Análise de dados (17/9/2012 e 18/9/2012): auditoria de tráfego**

Ficou acertado um período ‘oficial’ de 2 dias para aquisição simultânea de tráfego para *todos os pórticos*. Durante esse intervalo de tempo, foi feita também aquisição dos dados de tráfego, o que permitiu determinar o índice de perda de tráfego para cada sítio e período de 1 hora (que era o período mínimo de aferição ajustado pelos contadores). A determinação desse índice de perda é importante para se ponderar corretamente as atribuições de desempenho de cada sítio. Aquele ponto onde a taxa de perda é superior a um determinado valor deve ser reinterpretado em seu desempenho, já que perdas de tráfego devem ser mínimas (isto é, o valor de %fII deve ser o menor possível).

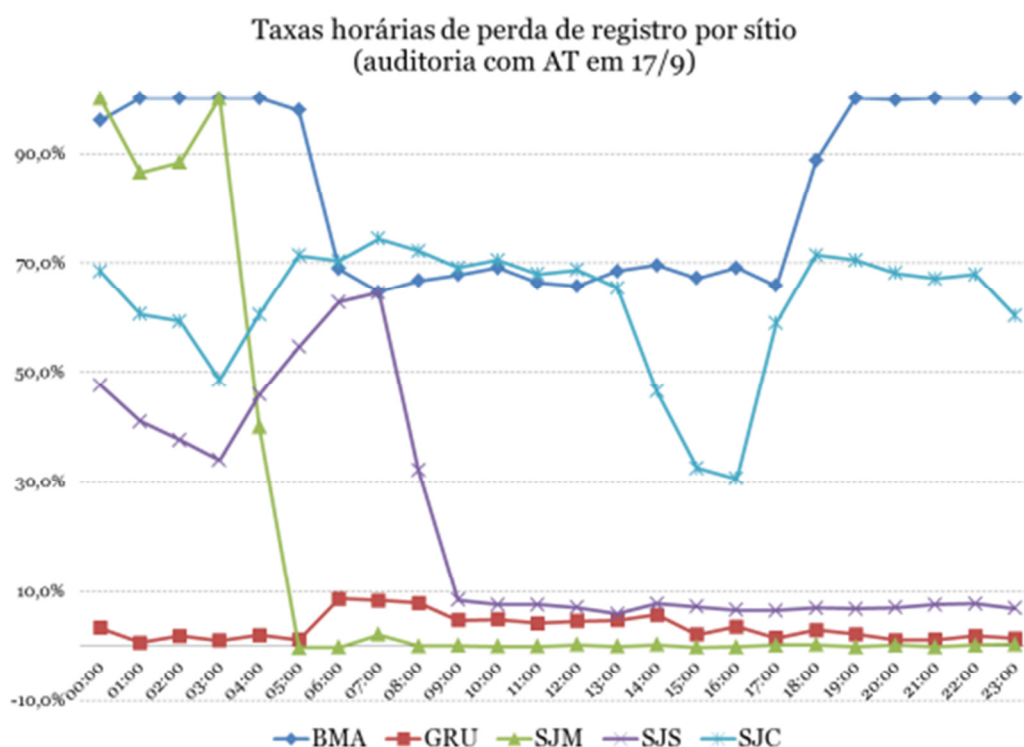


Figura 35 %IrV (taxa horária de perda de registro) por comparação com dados de AT no dia 17/9 por sítio.

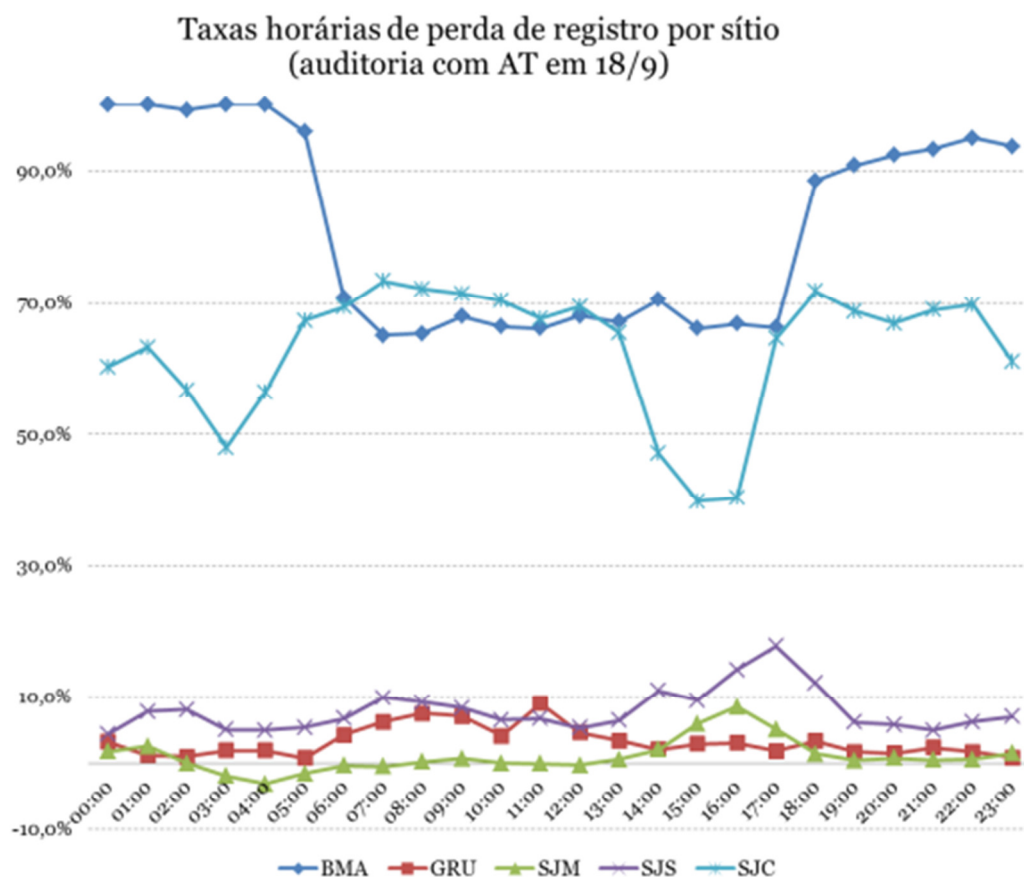




Figura 36 %IrV (taxa horária de perda de registro) por comparação com dados de AT no dia 18/9 por sítio.

Um resumo da aferição do tráfego por sítio pode ser visto na Figura 35 e Figura 36 onde valores de %IrV são plotados em cada hora do sistema. As curvas para BMA e SJC trazem valores elevados de perda de aquisição de tráfego (superiores em média a 50%). O fornecedor responsável pelo sítio de BMA informou que as perdas chegam a 100% no período noturno por conta do sistema de detecção de veículo usado que é fortemente dependente do grau de iluminação externo. Durante a noite, por não existir iluminação, o sistema de BMA simplesmente ‘para de funcionar’<sup>33</sup>. O valor aferido de perda para SJM foi relativamente elevado nas primeiras horas do dia 17/9 e, depois, estabilizou em valores baixo em todo o dia 17/9, enquanto que oscilou com valores negativos (excesso de registros) nas primeiras horas do dia 18/9. Situação semelhante ocorreu em SJS até as 9:00 do dia 17/9, apresentando um pico acima de 10% em torno das 17:00 do dia 18/9. Tais dados permitem compilar um resumo das frações de perda de registro por sítio conforme o gráfico da Figura 37. É possível ver que a taxa de perda de registro inferior a 10% ocorre com vários sistemas (GRU, SJM e SJS) como é esperado. Dado as assimetrias de volume de tráfego envolvidas (ver Figura 4), o melhor resultado da auditoria fica por conta do sítio de GRU que, além disso, apresentou uniformidade do índice por todos os períodos durante os dois dias.

---

<sup>33</sup> O fornecedor nos disse, entretanto, que essa limitação é de configuração de equipamento. Pelo uso de um sistema de laços reais (contadores) seria possível fazer o sistema capturar todos os veículos.

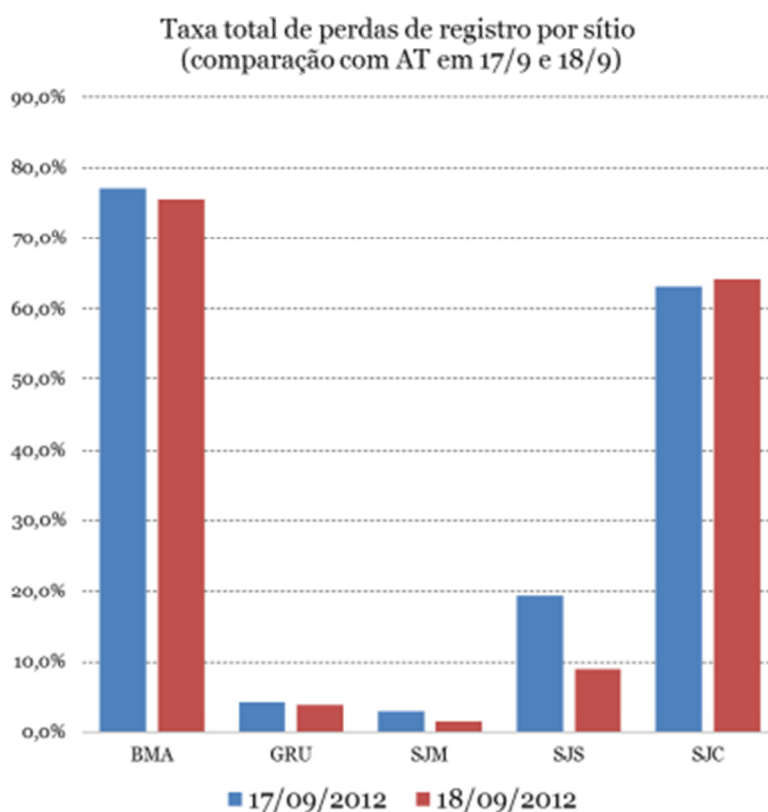


Figura 37 %IrV médio integrados por sítio (proporcional à perda de tráfego). O valor pequeno de SJM está contaminado pela presença de registros espúrios que contribuem para tornar %IrV negativo em vários períodos.

As seguintes tabelas (Figura 38 a Figura 47) trazem os relatórios de conferência de tráfego e registro. As colunas denominadas de ‘P’ trazem os resultados do tráfego totalizado para a pista indicada, enquanto que ‘RP’ são os registros de cada pista. A coluna TOTAL<R> traz o total de registros integrado para todas as pistas, enquanto que TOTAL traz a contagem ou registro por pista em cada sítio.

BMA	P1	P2	TOTAL	RP1	RP2	TOTAL<R>	IRV
00:00	130	291	421	17	0	17	96,0%
01:00	68	206	274	0	0	0	100,0%
02:00	39	116	155	0	0	0	100,0%
03:00	25	84	109	0	0	0	100,0%
04:00	44	143	187	0	0	0	100,0%
05:00	142	267	409	9	0	9	97,8%
06:00	382	567	949	295	0	295	68,9%
07:00	639	698	1337	472	0	472	64,7%
08:00	512	707	1219	405	0	405	66,8%
09:00	453	635	1088	350	0	350	67,8%
10:00	468	656	1124	347	0	347	69,1%
11:00	474	649	1123	378	0	378	66,3%
12:00	496	625	1121	384	0	384	65,7%
13:00	473	642	1115	351	0	351	68,5%
14:00	449	696	1145	348	0	348	69,6%
15:00	512	692	1204	395	0	395	67,2%
16:00	425	663	1088	336	0	336	69,1%
17:00	556	679	1235	422	0	422	65,8%
18:00	553	698	1251	141	0	141	88,7%
19:00	404	551	955	0	0	0	100,0%
20:00	283	490	773	2	0	2	99,7%
21:00	336	485	821	0	0	0	100,0%
22:00	258	472	730	0	0	0	100,0%
23:00	144	353	497	0	0	0	100,0%
<b>TOTAL</b>	<b>8265</b>	<b>12065</b>	<b>20330</b>	<b>4652</b>	<b>0</b>	<b>4652</b>	<b>77,1%</b>

Figura 38 Tabela com tráfego e registro de BMA em 17/9.

BMA	P1	P2	TOTAL	RP1	RP2	TOTAL<R>	IRV
00:00	78	276	354	0	0	0	100,0%
01:00	96	245	341	0	0	0	100,0%
02:00	51	200	251	2	0	2	99,2%
03:00	32	130	162	0	0	0	100,0%
04:00	24	155	179	0	0	0	100,0%
05:00	93	271	364	15	0	15	95,9%
06:00	327	557	884	258	0	258	70,8%
07:00	598	706	1304	456	0	456	65,0%
08:00	552	673	1225	425	0	425	65,3%
09:00	514	683	1197	384	0	384	67,9%
10:00	452	617	1069	359	0	359	66,4%
11:00	477	645	1122	380	0	380	66,1%
12:00	452	639	1091	349	0	349	68,0%
13:00	474	648	1122	369	0	369	67,1%
14:00	518	700	1218	360	0	360	70,4%
15:00	499	661	1160	393	0	393	66,1%
16:00	531	722	1253	416	0	416	66,8%
17:00	642	728	1370	462	0	462	66,3%
18:00	601	705	1306	151	0	151	88,4%
19:00	389	552	941	87	0	87	90,8%
20:00	283	498	781	60	0	60	92,3%
21:00	302	498	800	54	0	54	93,3%
22:00	216	418	634	32	0	32	95,0%
23:00	167	355	522	33	0	33	93,7%
<b>TOTAL</b>	<b>8368</b>	<b>12282</b>	<b>20650</b>	<b>5045</b>	<b>0</b>	<b>5045</b>	<b>75,6%</b>

Figura 39 Tabela com tráfego e registro de BMA em 18/9.

GRU	17/set	P1	P2	P3	P4	P5	P6	P7	P8	TOTAL	RP1	RP2	RP3	RP4	RP5	RP6	RP7	RP8	TOTAL <R>	IRV
	00:00	177	306	283	482	395	386	141	2	2172	170	303	273	471	376	373	132	2	2100	3,3%
	01:00	111	178	188	325	253	200	73	4	1332	109	176	188	325	251	198	73	4	1324	0,6%
	02:00	63	170	140	265	210	174	61	4	1087	63	171	139	257	208	167	58	4	1067	1,8%
	03:00	99	276	120	271	300	251	77	1	1395	97	271	122	272	292	250	76	1	1381	1,0%
	04:00	320	472	284	544	562	465	155	20	2822	310	470	282	535	548	453	148	19	2765	2,0%
	05:00	849	723	1269	1265	1202	1062	416	89	6875	838	711	1257	1243	1195	1054	411	89	6798	1,1%
	06:00	1737	1224	1723	1735	1867	1739	1352	225	11602	1540	1064	1514	1502	1805	1664	1300	217	10606	8,6%
	07:00	1837	1454	1531	1782	1861	1822	1401	188	11876	1640	1273	1363	1606	1770	1726	1337	179	10894	8,3%
	08:00	1392	955	1588	1767	1862	1648	1070	6	10288	1219	839	1434	1579	1789	1581	1034	6	9481	7,8%
	09:00	1261	899	1459	1546	1551	1360	798	32	8906	1177	845	1357	1436	1528	1333	776	34	8486	4,7%
	10:00	1107	738	720	700	758	886	1078	296	6283	1055	699	678	658	722	847	1026	295	5980	4,8%
	11:00	903	722	1188	1179	1199	1158	962	210	7521	877	693	1135	1124	1153	1110	916	199	7207	4,2%
	12:00	742	686	1324	1413	1462	1314	677	39	7657	703	641	1211	1319	1443	1289	663	40	7309	4,5%
	13:00	793	736	1315	1542	1488	1341	708	55	7978	733	699	1254	1455	1439	1286	685	56	7607	4,7%
	14:00	814	751	1330	1522	1449	1325	726	26	7943	741	681	1211	1403	1424	1294	716	23	7493	5,7%
	15:00	830	817	1280	1411	1407	1333	679	28	7785	816	798	1257	1383	1382	1305	657	28	7626	2,0%
	16:00	909	764	1311	1381	1297	1223	685	27	7597	857	723	1252	1322	1279	1196	672	27	7328	3,5%
	17:00	892	792	1396	1414	1350	1322	748	30	7944	874	780	1380	1385	1334	1307	739	28	7827	1,5%
	18:00	1019	825	1393	1337	1324	1332	678	36	7944	978	788	1338	1290	1307	1310	669	35	7715	2,9%
	19:00	569	565	1170	1134	1103	1109	542	17	6209	555	552	1132	1099	1095	1095	531	17	6076	2,1%
	20:00	388	427	884	921	854	806	335	20	4585	387	418	828	910	846	794	335	19	4537	1,0%
	21:00	413	420	700	898	795	732	317	31	4306	407	416	692	888	788	722	313	31	4257	1,1%
	22:00	395	493	709	887	824	714	292	17	4331	379	481	703	881	803	704	285	16	4252	1,8%
	23:00	241	351	373	589	429	167	1	1	2623	239	344	369	583	465	423	163	1	2587	1,4%
<b>TOTAL</b>		<b>17861</b>	<b>15744</b>	<b>23628</b>	<b>26310</b>	<b>25845</b>	<b>24131</b>	<b>14138</b>	<b>1404</b>	<b>149061</b>	<b>16764</b>	<b>14836</b>	<b>22369</b>	<b>24926</b>	<b>25242</b>	<b>23481</b>	<b>13715</b>	<b>1370</b>	<b>142703</b>	<b>4,3%</b>

Figura 40 Tabela com tráfego e registro de GRU em 17/9.

GRU	18/ago	P1	P2	P3	P4	P5	P6	P7	P8	TOTAL	RP1	RP2	RP3	RP4	RP5	RP6	RP7	RP8	TOTAL <R>	IRV
	00:00	123	254	187	357	309	289	67	0	1586	117	243	181	348	302	279	65		<b>1535</b>	3,2%
	01:00	93	234	137	274	239	184	68	1	1230	93	231	137	272	236	178	67	1	<b>1215</b>	1,2%
	02:00	91	236	104	230	265	207	67	0	1200	89	237	103	229	260	205	65		<b>1188</b>	1,0%
	03:00	116	311	107	264	329	305	99	0	1531	115	306	107	260	321	298	94		<b>1501</b>	2,0%
	04:00	290	456	288	545	584	494	183	23	2863	280	453	282	540	574	480	176	23	<b>2808</b>	1,9%
	05:00	694	649	1124	1181	1160	1051	439	88	6386	700	646	1114	1175	1146	1043	424	88	<b>6336</b>	0,8%
	06:00	1534	1160	1803	1823	1940	1816	1333	189	11598	1475	1107	1727	1757	1893	1766	1308	71	<b>11104</b>	4,3%
	07:00	1669	1319	1610	1739	1713	1662	1528	299	11539	1566	1235	1517	1639	1649	1605	1458	153	<b>10822</b>	6,2%
	08:00	1669	1089	1649	1815	1821	1684	965	60	10752	1487	980	1451	1602	1786	1646	938	58	<b>9948</b>	7,5%
	09:00	1342	968	1524	1628	1542	1393	890	73	9360	1299	946	1489	1588	1364	1231	732	45	<b>8694</b>	7,1%
	10:00	986	821	1000	925	940	994	998	303	6967	942	787	938	868	906	966	975	297	<b>6679</b>	4,1%
	11:00	1069	810	861	829	760	828	1065	394	6616	930	692	713	687	753	814	1048	383	<b>6020</b>	9,0%
	12:00	872	774	1205	1293	1240	1241	945	148	7718	814	719	1119	1194	1227	1219	922	146	<b>7360</b>	4,6%
	13:00	896	747	1362	1561	1510	1373	716	58	8223	865	705	1303	1494	1486	1339	694	55	<b>7941</b>	3,4%
	14:00	881	773	1280	1457	1449	1380	700	25	7945	845	761	1268	1425	1411	1358	685	25	<b>7778</b>	2,1%
	15:00	1041	819	1294	1426	1418	1329	738	16	8081	992	786	1244	1382	1398	1304	721	16	<b>7843</b>	2,9%
	16:00	859	780	1355	1437	1333	1304	681	37	7786	822	751	1300	1378	1312	1279	668	37	<b>7547</b>	3,1%
	17:00	930	767	1347	1367	1316	1284	738	38	7787	912	748	1323	1340	1299	1258	728	35	<b>7643</b>	1,8%
	18:00	935	779	1339	1318	1329	1303	721	30	7754	887	734	1287	1253	1314	1280	713	29	<b>7497</b>	3,3%
	19:00	604	570	1244	1178	1204	1121	543	20	6484	593	562	1221	1162	1182	1103	527	20	<b>6370</b>	1,8%
	20:00	383	465	923	1010	855	852	370	11	4869	377	461	907	991	843	837	367	11	<b>4794</b>	1,5%
	21:00	363	433	751	916	786	726	305	34	4314	353	422	720	883	781	719	300	34	<b>4212</b>	2,4%
	22:00	395	531	653	860	762	734	285	9	4229	389	516	641	850	752	722	278	9	<b>4157</b>	1,7%
	23:00	199	362	379	577	499	476	186	0	2678	201	361	374	575	496	471	176	0	<b>2654</b>	0,9%
<b>TOTAL</b>		<b>18034</b>	<b>16107</b>	<b>23526</b>	<b>26010</b>	<b>25303</b>	<b>24030</b>	<b>14630</b>	<b>1856</b>	<b>149496</b>	<b>17143</b>	<b>15389</b>	<b>22466</b>	<b>24892</b>	<b>24691</b>	<b>23400</b>	<b>14129</b>	<b>1536</b>	<b>143646</b>	<b>3,9%</b>

Figura 41 Tabela com tráfego e registro de GRU em 18/9.

SJM	17/set	P1	P2	P3	P4	P5	P6	P7	TOTAL	RP1	RP2	RP3	RP4	RP5	RP6	RP7	TOTAL <R>	IRV	
	00:00	106	196	0	294	336	1	1	934								0	100,0%	
	01:00	55	114	0	139	188	28	6	530	5	14			21	22	5	4	71	86,6%
	02:00	15	52	0	2	11	200	66	346					2	3	33	2	40	88,4%
	03:00	0	0	1	99	157	29	25	311									0	100,0%
	04:00	4	4	1	224	264	50	0	547	2	4		1	136	143	42		328	40,0%
	05:00	0	0	0	581	514	112	18	1225					583	517	111	18	1229	-0,3%
	06:00	0	0	0	1498	1052	316	54	2920					1498	1056	318	55	2927	-0,2%
	07:00	89	77	0	1676	1405	580	58	3885	89				1674	1404	579	57	3803	2,1%
	08:00	510	524	0	1274	1152	467	53	3980	511	524			1270	1153	468	54	3980	0,0%
	09:00	453	510	1	1434	1250	516	51	4215	454	509			1434	1249	517	50	4213	0,0%
	10:00	512	544	1	1517	1207	523	47	4351	511	546		1	1518	1209	522	47	4354	-0,1%
	11:00	509	535	0	1561	1260	458	47	4370	509	533			1560	1264	460	47	4373	-0,1%
	12:00	466	519	0	1748	1248	547	45	4573	468	520			1745	1242	544	45	4564	0,2%
	13:00	428	467	1	1461	1319	532	56	4264	429	467		1	1459	1321	535	56	4268	-0,1%
	14:00	534	533	1	1279	1233	956	127	4663	532	532		1	1274	1228	956	129	4652	0,2%
	15:00	494	505	0	686	597	1350	435	4067	493	506			689	600	1352	436	4076	-0,2%
	16:00	475	469	0	1174	1292	1303	109	4822	478	469			1175	1297	1303	107	4829	-0,1%
	17:00	436	370	0	1608	1435	781	60	4690	433	369			1609	1435	779	59	4684	0,1%
	18:00	511	379	1	1582	1439	711	67	4600	513	378			1574	1431	716	67	4679	0,2%
	19:00	355	360	0	627	654	1310	257	3563	352	361		1	631	658	1307	259	3569	-0,2%
	20:00	514	465	2	1592	1422	734	72	4801	515	466		2	1588	1419	732	71	4793	0,2%
	21:00	375	365	0	1191	1010	402	43	3386	375	365			1193	1012	402	43	3390	-0,1%
	22:00	353	406	0	1097	994	425	23	3298	354	405			1094	991	426	22	3292	0,2%
	23:00	188	296	0	510	566	165	21	1746	187	295			509	565	164	21	1741	0,3%
total		7382	7600	0	24854	22005	12406	1741	76177	7210	7263		7	24236	21210	12271	1640	73855	3,0%

SJM	18/set	P1	P2	P3	P4	P5	P6	P7	TOTAL	RP1	RP2	RP3	RP4	RP5	RP6	RP7	TOTAL <R>	IRV
	00:00	106	213	0	217	325	62	5	1500	107	213		215	325	61	5	1473	1,8%
	01:00	6	24	0	227	277	36	2	984	5			226	278	36	2	959	2,5%
	02:00	0	1	0	140	234	34	3	789		1		140	234	34	3	789	0,0%
	03:00	0	0	0	130	213	34	0	972				130	212	34	1	990	-1,9%
	04:00	2	3	0	250	282	57	1	1871	2	3		251	283	57	17	1929	-3,1%
	05:00	0	0	0	612	536	111	17	4169				616	538	112	50	4233	-1,5%
	06:00	0	0	0	1478	1034	331	50	6909				1481	1040	331	65	6930	-0,3%
	07:00	56	64	0	1678	1423	730	65	8431	57	66		1673	1419	734	64	8467	-0,4%
	08:00	548	538	0	1364	1268	633	64	8672	548	538		1366	1267	631	104	8650	0,3%
	09:00	542	485	34	990	1051	1051	104	8524	542	483	34	984	1050	1049	54	8460	0,8%
	10:00	437	494	60	1388	1229	605	54	8551	442	497	60	1387	1228	607	43	8546	0,1%
	11:00	468	494	0	1531	1246	502	43	8656	463	494		1532	1247	502	44	8657	0,0%
	12:00	455	492	0	1712	1186	483	44	8738	455	494		1707	1187	483	49	8760	-0,3%
	13:00	426	463	0	1715	1181	532	49	9078	427	461		1717	1182	533	65	9023	0,6%
	14:00	453	474	0	1766	1417	538	64	9537	447	463		1739	1398	532	59	9336	2,1%
	15:00	491	481	2	1644	1419	725	63	9747	479	463	2	1604	1386	707	57	9164	6,0%
	16:00	493	483	2	1666	1436	774	68	9668	451	434	2	1513	1307	704	55	8848	8,5%
	17:00	445	408	2	1632	1436	761	62	9602	408	374	2	1510	1331	697	60	9110	5,1%
	18:00	630	518	6	1534	1415	689	64	9502	614	503	6	1496	1372	666	71	9365	1,4%
	19:00	521	435	0	1581	1401	637	71	8771	517	436		1582	1405	637	60	8733	0,4%
	20:00	432	404	1	1457	1240	531	60	7631	431	400	1	1454	1234	528	48	7571	0,8%
	21:00	331	367	1	1191	1068	500	48	6628	332	367	1	1187	1066	501	21	6594	0,5%
	22:00	331	366	0	996	932	476	21	4924	331	368		993	936	477	14	4894	0,6%
	23:00	212	310	1	542	540	183	14	80094	213	308	1	537	534	182		78866	1,5%
total		7385	7517	109	27441	23789	11015	1036	233948	7271	7366	109	27040	23459	10835	1011	230347	1,5%

Figura 43 Tabela com tráfego e registro de SJM em 18/9. Valores negativos de Irv correspondem a registros espúrios gerados pelo sistema. Observe que em SJM os valores negativos compensam, na média, os positivos reduzindo o IrV médio.

SJC	17/set	P1	P2	total	RP1	RP2	TOTAL<r>	IRV
	00:00	236	293	529	0	277	277	47,6%
	01:00	155	240	395	0	233	233	41,0%
	02:00	153	312	465	0	290	290	37,6%
	03:00	138	309	447	0	295	295	34,0%
	04:00	290	384	674	0	365	365	45,8%
	05:00	601	573	1174	0	533	533	54,6%
	06:00	1348	901	2249	0	836	836	62,8%
	07:00	1936	1179	3115	0	1104	1104	64,6%
	08:00	1579	994	2573	820	925	1745	32,2%
	09:00	1359	933	2292	1223	877	2100	8,4%
	10:00	1223	848	2071	1106	809	1915	7,5%
	11:00	1199	857	2056	1084	817	1901	7,5%
	12:00	1156	832	1988	1053	796	1849	7,0%
	13:00	1307	879	2186	1207	852	2059	5,8%
	14:00	1331	926	2257	1205	878	2083	7,7%
	15:00	1577	974	2551	1455	914	2369	7,1%
	16:00	1484	930	2414	1360	897	2257	6,5%
	17:00	1962	1169	3131	1788	1141	2929	6,5%
	18:00	1777	1041	2818	1638	986	2624	6,9%
	19:00	1271	822	2093	1167	785	1952	6,7%
	20:00	764	631	1395	697	600	1297	7,0%
	21:00	580	547	1127	527	515	1042	7,5%
	22:00	468	493	961	421	466	887	7,7%
	23:00	375	441	816	342	418	760	6,9%
TOTAL		24269	17508	41777	17093	16609	33702	19,3%

Figura 44 Tabela com tráfego e registro de SJS em 17/9.

SJC	18/set	P1	P2	total	RP1	RP2	TOTAL<R>	IRV
	00:00	199	326	525	186	316	502	4,4%
	01:00	154	320	474	140	297	437	7,8%
	02:00	151	343	494	139	315	454	8,1%
	03:00	140	319	459	122	314	436	5,0%
	04:00	212	354	566	201	337	538	4,9%
	05:00	480	561	1041	447	538	985	5,4%
	06:00	1227	891	2118	1134	842	1976	6,7%
	07:00	1871	1182	3053	1645	1105	2750	9,9%
	08:00	1589	997	2586	1420	928	2348	9,2%
	09:00	1401	961	2362	1262	903	2165	8,3%
	10:00	1274	861	2135	1174	821	1995	6,6%
	11:00	1120	817	1937	1020	787	1807	6,7%
	12:00	1271	877	2148	1183	850	2033	5,4%
	13:00	1320	905	2225	1222	860	2082	6,4%
	14:00	1358	954	2312	1162	896	2058	11,0%
	15:00	1579	998	2577	1443	889	2332	9,5%
	16:00	1636	978	2614	1359	886	2245	14,1%
	17:00	2009	1202	3211	1510	1132	2642	17,7%
	18:00	1776	1054	2830	1490	998	2488	12,1%
	19:00	1248	802	2050	1147	777	1924	6,1%
	20:00	827	651	1478	766	626	1392	5,8%
	21:00	617	568	1185	575	551	1126	5,0%
	22:00	552	517	1069	506	496	1002	6,3%
	23:00	371	445	816	336	423	759	7,0%
<b>TOTAL</b>		<b>24382</b>	<b>17883</b>	<b>42265</b>	<b>21589</b>	<b>16887</b>	<b>38476</b>	<b>9,0%</b>

Figura 45 Tabela com tráfego e registro de SJS em 18/9.

SJC	17/set	P1	P2	P3	TOTAL	RP1	RP2	RP3	TOTAL<R>	IRV
	00:00	220	289	143	652	61	84	60	205	68,6%
	01:00	109	191	104	404	50	82	27	159	60,6%
	02:00	50	152	73	275	31	68	13	112	59,3%
	03:00	47	169	89	305	51	87	19	157	48,5%
	04:00	99	231	159	489	67	91	35	193	60,5%
	05:00	482	527	301	1310	108	165	102	375	71,4%
	06:00	1332	1020	529	2881	162	335	357	854	70,4%
	07:00	2071	1370	791	4232	212	365	504	1081	74,5%
	08:00	1699	1147	613	3459	202	334	426	962	72,2%
	09:00	1319	1023	498	2840	185	338	354	877	69,1%
	10:00	1124	927	478	2529	184	259	303	746	70,5%
	11:00	1120	931	497	2548	201	297	317	815	68,0%
	12:00	1125	900	507	2532	170	291	330	791	68,8%
	13:00	1283	960	535	2778	185	337	437	959	65,5%
	14:00	1227	999	545	2771	241	441	802	1484	46,4%
	15:00	1262	954	538	2754	288	566	1003	1857	32,6%
	16:00	1354	950	620	2924	335	616	1076	2027	30,7%
	17:00	1733	1195	648	3576	230	582	654	1466	59,0%
	18:00	1647	1137	627	3411	194	328	452	974	71,4%
	19:00	1245	938	462	2645	153	286	341	780	70,5%
	20:00	786	688	386	1860	142	226	224	592	68,2%
	21:00	610	610	335	1555	128	193	190	511	67,1%
	22:00	534	587	327	1448	122	199	144	465	67,9%
	23:00	369	521	319	1209	133	215	131	479	60,4%
<b>TOTAL</b>		<b>22847</b>	<b>18416</b>	<b>10124</b>	<b>51387</b>	<b>3835</b>	<b>6785</b>	<b>8301</b>	<b>18921</b>	<b>63,2%</b>

Figura 46 Tabela com tráfego e registro de SJC em 17/9.

SJC	18/set	P1	P2	P3	TOTAL	RP1	RP2	RP3	TOTAL<R>	IRV
	00:00	198	366	226	<b>790</b>	94	164	56	<b>314</b>	60,3%
	01:00	97	245	153	<b>495</b>	63	94	25	<b>182</b>	63,2%
	02:00	51	160	94	<b>305</b>	43	70	19	<b>132</b>	56,7%
	03:00	42	180	122	<b>344</b>	63	102	14	<b>179</b>	48,0%
	04:00	90	246	151	<b>487</b>	69	122	21	<b>212</b>	56,5%
	05:00	447	556	314	<b>1317</b>	113	212	106	<b>431</b>	67,3%
	06:00	1207	983	533	<b>2723</b>	181	334	319	<b>834</b>	69,4%
	07:00	2009	1345	791	<b>4145</b>	241	380	487	<b>1108</b>	73,3%
	08:00	1713	1217	631	<b>3561</b>	195	347	451	<b>993</b>	72,1%
	09:00	1366	1010	481	<b>2857</b>	159	313	345	<b>817</b>	71,4%
	10:00	1102	937	517	<b>2556</b>	196	266	297	<b>759</b>	70,3%
	11:00	1098	918	471	<b>2487</b>	168	328	310	<b>806</b>	67,6%
	12:00	1099	890	475	<b>2464</b>	164	286	304	<b>754</b>	69,4%
	13:00	1273	971	565	<b>2809</b>	208	326	436	<b>970</b>	65,5%
	14:00	1240	966	569	<b>2775</b>	274	432	763	<b>1469</b>	47,1%
	15:00	1242	988	576	<b>2806</b>	290	603	796	<b>1689</b>	39,8%
	16:00	1353	1060	618	<b>3031</b>	308	567	934	<b>1809</b>	40,3%
	17:00	1771	1185	660	<b>3616</b>	244	447	591	<b>1282</b>	64,5%
	18:00	1733	1175	662	<b>3570</b>	170	350	488	<b>1008</b>	71,8%
	19:00	1251	922	487	<b>2660</b>	184	300	348	<b>832</b>	68,7%
	20:00	775	688	383	<b>1846</b>	143	235	234	<b>612</b>	66,8%
	21:00	719	688	365	<b>1772</b>	122	225	204	<b>551</b>	68,9%
	22:00	564	645	323	<b>1532</b>	123	197	144	<b>464</b>	69,7%
	23:00	363	524	280	<b>1167</b>	111	214	129	<b>454</b>	61,1%
<b>TOTAL</b>		<b>22803</b>	<b>18865</b>	<b>10447</b>	<b>52115</b>	<b>3926</b>	<b>6914</b>	<b>7821</b>	<b>18661</b>	<b>64,2%</b>

Figura 47 Tabela com tráfego e registro de SJC em 18/9.

### Tempos de inspeção em BackOffice

Esta seção apresenta as médias de tempo de inspeção por tipo de registro, conforme a classificação utilizada segundo a seção “Hierarquia binomial usada e notas finais”. Os tempos de inspeção foram medidos em milissegundos de forma automática à medida que o inspetor executava a operação em BackOffice, através de um relógio interno da interface.

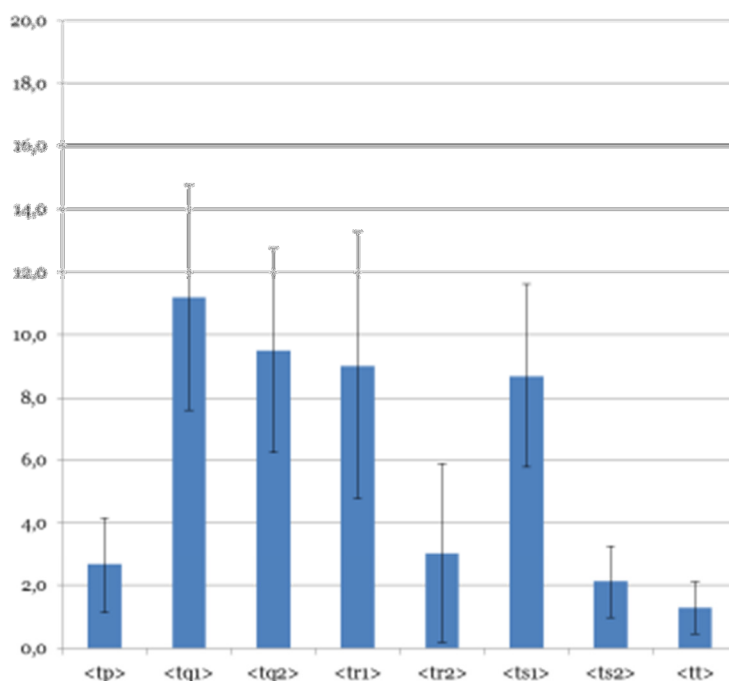


Figura 48 Tempo de inspeção do Inspetor 1 para as várias classes. Dados de 6/9, 7/9 e 8/9.



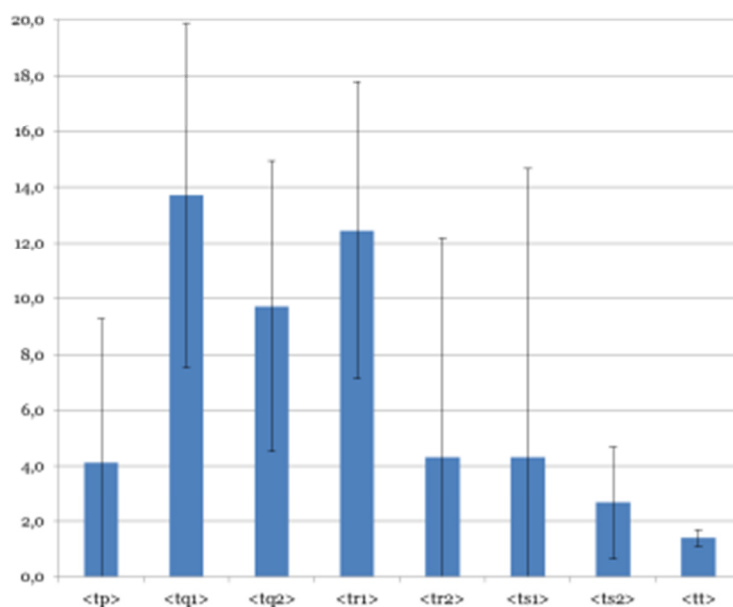


Figura 49 Tempo de inspeção do Inspetor 2 para as várias classes. Dados de 17/9 e 18/9.

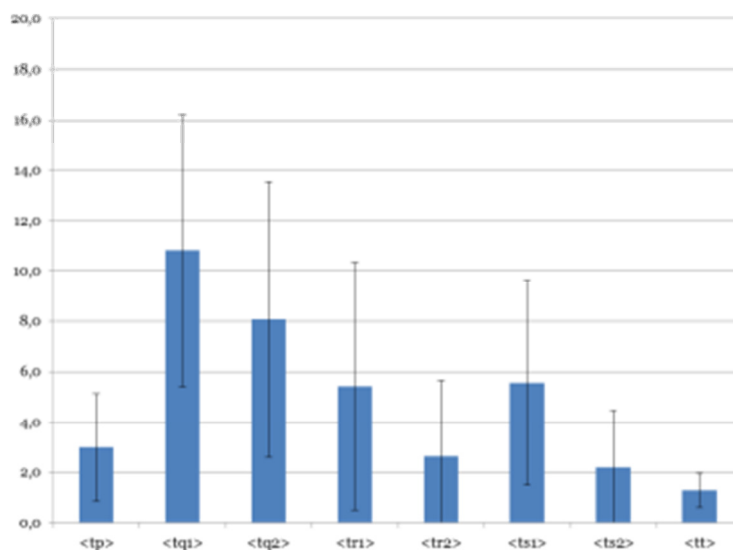


Figura 50 Tempo de inspeção do Inspetor 3 para as várias classes. Dados de 17/9 e 18/9.

A Figura 49 e Figura 50 trazem as médias e desvios padrão para três inspetores. A sequência de tempos (abcissa) é chamado de <tp>, <tq1>, <tq2>, <tr1>, <tr2>, <ts1>, <ts2> e <tt> ou tempos de inspeção de imagens corretamente validadas pelo software OCR, imagens com código válido, mas só legíveis por humanos, imagens sem código válido legíveis por humanos, imagens com código válido ilegíveis por humanos, imagens sem código válido ilegíveis por humanos, imagens com código válido sem placa, imagens sem código válido sem placa e imagens contendo motos, respectivamente. Para a compilação desses resultados, foram utilizados dados dos dias 6/9, 7/9 e 8/9 e a série final de 17/9 e 18/9.

Os tempos médios ponderados pelos totais de cada classe foram:

- Inspetor 1: 3,4 s;
- Inspetor 2: 5,0 s;
- Inspetor 3: 2,8 s.

O que demonstra haver uma variação de até 40% nos tempos de inspeção para esses inspetores. O tempo máximo de inspeção foi quase 14 s para a classe q1 e inspetor 2 e o tempo mínimo foi de 2,1 s para a classe p do inspetor 3. As seguintes observações são pertinentes para o comportamento observado dos tempos:

- Para todos os inspetores, foi observado que as classes p, r2, s2 e t são as que consomem os menores tempos, pois o operador precisa apenas clicar no botão ‘OK’ para validar o resultado (seja porque o código fornecido é o correto, a imagem não contém uma placa (s2) ou o inspetor não consegue ler (r2) um registro que não tem código associado);
- A classe q1 é a que precisa de maior tempo, pois o inspetor tem que checar se o código fornecido é válido e digitar o código correto;
- A classe q2 também consome tempo, pois o inspetor tem que digitar o código, embora leve menos tempo que q1 porque não há necessidade de checagem;
- A classe r1 também exige maior tempo, pois o inspetor tem que checar se o código está correto e decidir se a imagem é ilegível para ele;

A média dos valores fornecidos para os tempos globais de inspeção é da ordem de 4 s. Uma vez que boa parte do tempo é gasto nas classes q e r1 (para as quais o inspetor tem que digitar o código ou decidir por considerar a imagem ilegível), está claro que a média dos tempos diminuirá caso um sistema retorne menos imagens ilegíveis ou com códigos inválidos. Como consequência, quanto melhor a eficiência do sistema de leitura automática, menor o tempo de inspeção em BackOffice e, portanto, menor é o esforço humano para validação dos registros.

### **Análise de dados: evolução das frações (17/9/2012 e 18/9/2012)**

Apresentamos aqui as tabelas da evolução das notas consolidadas por sítio e período horário para os dias 17/9 e 18/9. Ao invés de reproduzir o índice %IrV, optamos por exibir uma coluna com seu complemento ( $1-\%IrV$ ), de forma que o resultado mostrado é a quantidade de fluxo captado por cada um dos sistemas e é obtida pelo produto de ( $1-\%IrV$ ) com a respectiva nota.. Os resultados estão nas tabelas da Figura 51 a Figura 60.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sft	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	66,9%	9,9%	12,0%	10,6%	0,7%	92,2%	72,5%	3,9%	2,5%	2,7%	2,6%	0,7%	2,6%	3,7%	96,7%	64,7%	9,5%
1	63,9%	11,8%	14,6%	9,0%	0,7%	86,8%	73,6%	4,0%	2,7%	2,9%	2,4%	0,7%	3,3%	3,7%	99,4%	63,5%	11,7%
2	58,3%	13,2%	16,0%	11,1%	1,4%	91,3%	63,9%	4,1%	2,8%	3,1%	2,6%	1,0%	2,9%	4,0%	98,2%	57,3%	13,0%
3	62,4%	7,1%	15,6%	13,5%	1,4%	95,7%	65,2%	4,1%	2,2%	3,1%	2,9%	1,0%	2,1%	4,0%	99,0%	61,8%	7,0%
4	48,7%	14,0%	24,0%	12,7%	0,7%	84,9%	57,3%	4,1%	2,8%	3,5%	2,7%	0,7%	3,9%	4,0%	98,0%	47,7%	13,7%
5	63,8%	8,1%	11,9%	10,0%	6,3%	94,4%	67,5%	3,8%	2,2%	2,6%	2,4%	1,9%	2,2%	3,7%	98,9%	63,0%	8,0%
6	70,0%	10,6%	7,5%	9,4%	2,5%	90,3%	77,5%	3,6%	2,4%	2,1%	2,3%	1,2%	2,7%	3,3%	91,4%	64,0%	9,7%
7	71,3%	6,9%	6,3%	10,6%	5,0%	91,9%	77,5%	3,6%	2,0%	1,9%	2,4%	1,7%	2,4%	3,3%	91,7%	65,4%	6,3%
8	78,1%	5,5%	2,7%	8,9%	4,8%	98,3%	79,5%	3,4%	1,9%	1,4%	2,4%	1,8%	1,2%	3,3%	92,2%	72,0%	5,0%
9	71,6%	6,5%	5,8%	5,8%	10,3%	95,7%	74,8%	3,6%	2,0%	1,9%	1,9%	2,4%	1,9%	3,5%	95,3%	68,2%	6,1%
10	73,1%	7,5%	8,1%	9,4%	1,9%	95,1%	76,9%	3,5%	2,1%	2,2%	2,3%	1,1%	1,9%	3,3%	95,2%	69,6%	7,1%
11	74,4%	7,5%	6,3%	8,8%	3,1%	94,4%	78,8%	3,5%	2,1%	1,9%	2,2%	1,4%	2,0%	3,2%	95,8%	71,3%	7,2%
12	73,1%	6,9%	5,6%	7,5%	6,9%	94,4%	77,5%	3,5%	2,0%	1,8%	2,1%	2,0%	2,1%	3,3%	95,5%	69,8%	6,6%
13	75,6%	7,5%	8,1%	3,8%	5,0%	95,3%	79,4%	3,4%	2,1%	2,2%	1,5%	1,7%	1,9%	3,2%	95,3%	72,1%	7,2%
14	70,7%	8,3%	7,6%	5,7%	7,6%	93,3%	75,8%	3,6%	2,2%	2,1%	1,9%	2,1%	2,3%	3,4%	94,3%	66,7%	7,8%
15	67,5%	6,9%	11,3%	7,5%	6,9%	94,7%	71,3%	3,7%	2,0%	2,5%	2,1%	2,0%	2,1%	3,6%	98,0%	66,1%	6,7%
16	62,5%	9,4%	16,9%	6,3%	5,0%	92,6%	67,5%	3,8%	2,3%	3,0%	1,9%	1,7%	2,5%	3,7%	96,5%	60,3%	9,0%
17	65,6%	6,3%	14,4%	8,1%	5,6%	92,1%	71,3%	3,8%	1,9%	2,8%	2,2%	1,8%	2,5%	3,6%	98,5%	64,7%	6,2%
18	56,3%	11,9%	13,8%	15,0%	3,1%	90,9%	61,9%	3,9%	2,6%	2,7%	2,8%	1,4%	2,9%	3,8%	97,1%	54,6%	11,5%
19	55,1%	9,0%	16,7%	14,7%	4,5%	86,0%	64,1%	4,0%	2,3%	3,0%	2,8%	1,7%	3,5%	3,8%	97,9%	53,9%	8,8%
20	59,0%	12,2%	12,8%	13,5%	2,6%	86,0%	68,6%	3,9%	2,6%	2,7%	2,7%	1,3%	3,4%	3,7%	99,0%	58,4%	12,1%
21	56,3%	14,4%	15,6%	8,1%	5,6%	90,0%	62,5%	3,9%	2,8%	2,9%	2,2%	1,8%	3,0%	3,8%	98,9%	55,6%	14,2%
22	52,6%	14,7%	15,4%	12,8%	4,5%	84,5%	62,2%	4,0%	2,8%	2,9%	2,7%	1,7%	3,7%	3,9%	98,2%	51,6%	14,5%
23	69,5%	9,2%	8,5%	9,9%	2,8%	95,1%	73,0%	3,9%	2,4%	2,3%	2,5%	1,4%	2,1%	3,7%	98,6%	68,5%	9,1%

Figura 51 Tabela com os índices e erros associados para GRU 17/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sft	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	60,0%	7,9%	15,7%	15,7%	0,7%	88,4%	67,9%	4,1%	2,3%	3,1%	3,1%	0,7%	3,3%	3,9%	96,8%	58,1%	7,6%
1	60,3%	10,6%	17,7%	10,6%	0,7%	94,4%	63,8%	4,1%	2,6%	3,2%	2,6%	0,7%	2,4%	4,0%	98,8%	59,5%	10,5%
2	62,9%	10,7%	17,9%	7,9%	0,7%	87,1%	72,1%	4,1%	2,6%	3,2%	2,3%	0,7%	3,3%	3,8%	99,0%	62,2%	10,6%
3	55,0%	19,3%	10,7%	13,6%	1,4%	81,9%	67,1%	4,2%	3,3%	2,6%	2,9%	1,0%	4,0%	4,0%	98,0%	53,9%	18,9%
4	48,0%	19,7%	17,1%	12,5%	2,6%	86,9%	55,3%	4,1%	3,2%	3,1%	2,7%	1,3%	3,7%	4,0%	98,1%	47,1%	19,4%
5	60,6%	11,3%	10,0%	14,4%	3,8%	91,5%	66,3%	3,9%	2,5%	2,4%	2,8%	1,5%	2,7%	3,7%	99,2%	60,2%	11,2%
6	69,3%	10,7%	10,0%	6,7%	3,3%	88,9%	78,0%	3,8%	2,5%	2,4%	2,0%	1,5%	2,9%	3,4%	95,7%	66,4%	10,2%
7	71,9%	8,1%	6,9%	9,4%	3,8%	91,3%	78,8%	3,6%	2,2%	2,0%	2,3%	1,5%	2,5%	3,2%	93,8%	67,4%	7,6%
8	69,4%	5,0%	6,9%	6,3%	12,5%	94,1%	73,8%	3,6%	1,7%	2,0%	1,9%	2,6%	2,2%	3,5%	92,5%	64,2%	4,6%
9	73,8%	6,3%	4,4%	10,6%	5,0%	94,4%	78,1%	3,5%	1,9%	1,6%	2,4%	1,7%	2,1%	3,3%	92,9%	68,5%	5,8%
10	75,6%	8,8%	3,8%	11,3%	0,6%	95,3%	79,4%	3,4%	2,2%	1,5%	2,5%	0,6%	1,9%	3,2%	95,9%	72,5%	8,4%
11	78,8%	7,5%	3,8%	7,5%	2,5%	97,7%	80,6%	3,2%	2,1%	1,5%	2,1%	1,2%	1,3%	3,1%	91,0%	71,7%	6,8%
12	76,9%	4,4%	8,8%	6,9%	3,1%	96,1%	80,0%	3,3%	1,6%	2,2%	2,0%	1,4%	1,7%	3,2%	95,4%	73,3%	4,2%
13	76,3%	7,5%	5,0%	8,1%	3,1%	92,4%	82,5%	3,4%	2,1%	1,7%	2,2%	1,4%	2,3%	3,0%	96,6%	73,6%	7,2%
14	71,9%	8,8%	5,6%	6,9%	6,9%	94,3%	76,3%	3,6%	2,2%	1,8%	2,0%	2,0%	2,1%	3,4%	97,9%	70,4%	8,6%
15	63,9%	10,3%	11,0%	11,0%	3,9%	90,0%	71,0%	3,9%	2,4%	2,5%	2,5%	1,5%	2,9%	3,6%	97,1%	62,0%	10,0%
16	60,0%	8,1%	13,8%	11,3%	6,9%	92,3%	65,0%	3,9%	2,2%	2,7%	2,5%	2,0%	2,6%	3,8%	96,9%	58,2%	7,9%
17	68,8%	10,0%	8,8%	7,5%	5,0%	90,2%	76,3%	3,7%	2,4%	2,2%	2,1%	1,7%	2,7%	3,4%	98,2%	67,5%	9,8%
18	62,5%	8,1%	14,4%	7,5%	7,5%	89,3%	70,0%	3,8%	2,2%	2,8%	2,1%	2,1%	2,9%	3,6%	96,7%	60,4%	7,9%
19	54,4%	14,4%	18,8%	10,0%	2,5%	84,5%	64,4%	3,9%	2,8%	3,1%	2,4%	1,2%	3,6%	3,8%	98,2%	53,4%	14,1%
20	60,3%	9,9%	13,9%	11,3%	4,6%	90,1%	66,9%	4,0%	2,4%	2,8%	2,6%	1,7%	3,0%	3,8%	98,5%	59,3%	9,8%
21	57,9%	14,5%	8,2%	16,4%	3,1%	88,5%	65,4%	3,9%	2,8%	2,2%	2,9%	1,4%	3,1%	3,8%	97,6%	56,5%	14,1%
22	63,8%	12,8%	13,4%	9,4%	0,7%	88,8%	71,8%	3,9%	2,7%	2,8%	2,4%	0,7%	3,1%	3,7%	98,3%	62,7%	12,5%
23	57,9%	18,6%	13,6%	8,6%	1,4%	91,0%	63,6%	4,2%	3,3%	2,9%	2,4%	1,0%	3,0%	4,1%	99,1%	57,3%	18,4%

Figura 52 Tabela com os índices e erros associados para GRU 18/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	62,5%	7,5%	17,5%	12,5%	0,0%	78,1%	80,0%	7,7%	4,2%	6,0%	5,2%	0,0%	7,3%	6,3%	52,4%	32,7%	3,9%
1	65,0%	25,0%	5,0%	5,0%	0,0%	83,9%	77,5%	7,5%	6,8%	3,4%	3,4%	0,0%	6,6%	6,6%	59,0%	38,3%	14,7%
2	47,5%	25,0%	15,0%	10,0%	2,5%	76,0%	62,5%	7,9%	6,8%	5,6%	4,7%	2,5%	8,5%	7,7%	62,4%	29,6%	15,6%
3	55,0%	22,5%	17,5%	5,0%	0,0%	64,7%	85,0%	7,9%	6,6%	6,0%	3,4%	0,0%	8,2%	5,6%	66,0%	36,3%	14,8%
4	72,5%	5,0%	20,0%	2,5%	0,0%	90,6%	80,0%	7,1%	3,4%	6,3%	2,5%	0,0%	5,2%	6,3%	54,2%	39,3%	2,7%
5	77,5%	12,5%	2,5%	0,0%	7,5%	91,2%	85,0%	6,6%	5,2%	2,5%	0,0%	4,2%	4,9%	5,6%	45,4%	35,2%	5,7%
6	77,5%	7,5%	7,5%	2,5%	5,0%	88,6%	87,5%	6,6%	4,2%	4,2%	2,5%	3,4%	5,4%	5,2%	37,2%	28,8%	2,8%
7	90,0%	2,5%	0,0%	0,0%	7,5%	97,3%	92,5%	4,7%	2,5%	0,0%	0,0%	4,2%	2,7%	4,2%	35,4%	31,9%	0,9%
8	87,5%	5,0%	2,5%	3,8%	1,3%	95,9%	91,3%	3,7%	2,4%	1,7%	2,1%	1,2%	2,3%	3,2%	67,8%	59,3%	3,4%
9	82,5%	8,8%	0,0%	1,3%	7,5%	91,7%	90,0%	4,2%	3,2%	0,0%	1,2%	2,9%	3,3%	3,4%	91,6%	75,6%	8,0%
10	88,8%	2,5%	0,0%	6,3%	2,5%	95,9%	92,5%	3,5%	1,7%	0,0%	2,7%	1,7%	2,3%	2,9%	92,5%	82,1%	2,3%
11	81,3%	11,3%	2,5%	2,5%	2,5%	85,5%	95,0%	4,4%	3,5%	1,7%	1,7%	1,7%	4,0%	2,4%	92,5%	75,1%	10,4%
12	87,5%	3,8%	1,3%	0,0%	7,5%	94,6%	92,5%	3,7%	2,1%	1,2%	0,0%	2,9%	2,6%	2,9%	93,0%	81,4%	3,5%
13	91,3%	5,0%	3,8%	0,0%	0,0%	94,8%	96,3%	3,2%	2,4%	2,1%	0,0%	0,0%	2,5%	2,1%	94,2%	85,9%	4,7%
14	88,8%	3,8%	0,0%	3,8%	3,8%	93,4%	95,0%	3,5%	2,1%	0,0%	2,1%	2,1%	2,8%	2,4%	92,3%	81,9%	3,5%
15	87,5%	2,5%	5,0%	1,3%	3,8%	97,2%	90,0%	3,7%	1,7%	2,4%	1,2%	2,1%	1,9%	3,4%	92,9%	81,3%	2,3%
16	83,8%	3,8%	2,5%	6,3%	3,8%	93,1%	90,0%	4,1%	2,1%	1,7%	2,7%	2,1%	3,0%	3,4%	93,5%	78,3%	3,5%
17	83,8%	3,8%	7,5%	3,8%	1,3%	90,5%	92,5%	4,1%	2,1%	2,9%	2,1%	1,2%	3,4%	2,9%	93,5%	78,3%	3,5%
18	65,0%	11,3%	15,0%	1,3%	7,5%	82,5%	78,8%	5,3%	3,5%	4,0%	1,2%	2,9%	4,8%	4,6%	93,1%	60,5%	10,5%
19	61,3%	17,5%	12,5%	2,5%	6,3%	79,0%	77,5%	5,4%	4,2%	3,7%	1,7%	2,7%	5,2%	4,7%	93,3%	57,1%	16,3%
20	67,5%	15,0%	10,0%	6,3%	1,3%	87,1%	77,5%	5,2%	4,0%	3,4%	2,7%	1,2%	4,3%	4,7%	93,0%	62,8%	13,9%
21	63,8%	11,3%	16,3%	7,5%	1,3%	82,3%	77,5%	5,4%	3,5%	4,1%	2,9%	1,2%	4,9%	4,7%	92,5%	58,9%	10,4%
22	66,3%	17,5%	11,3%	0,0%	5,0%	82,8%	80,0%	5,3%	4,2%	3,5%	0,0%	2,4%	4,7%	4,5%	92,3%	61,1%	16,2%
23	73,8%	10,0%	12,5%	2,5%	1,3%	92,2%	80,0%	4,9%	3,4%	3,7%	1,7%	1,2%	3,4%	4,5%	93,1%	68,7%	9,3%

Figura 53 Tabela com os índices e erros associados para SJS 17/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	58,8%	12,5%	22,5%	5,0%	1,3%	81,0%	72,5%	5,5%	3,7%	4,7%	2,4%	1,2%	5,1%	5,0%	95,6%	56,2%	12,0%
1	76,3%	8,8%	7,5%	7,5%	0,0%	93,8%	81,3%	4,8%	3,2%	2,9%	2,9%	0,0%	3,0%	4,4%	92,2%	70,3%	8,1%
2	62,5%	8,8%	21,3%	7,5%	0,0%	87,7%	71,3%	5,4%	3,2%	4,6%	2,9%	0,0%	4,3%	5,1%	91,9%	57,4%	8,0%
3	62,5%	10,0%	22,5%	5,0%	0,0%	86,2%	72,5%	5,4%	3,4%	4,7%	2,4%	0,0%	4,5%	5,0%	95,0%	59,4%	9,5%
4	68,8%	11,3%	11,3%	7,5%	1,3%	88,7%	77,5%	5,2%	3,5%	3,5%	2,9%	1,2%	4,0%	4,7%	95,1%	65,3%	10,7%
5	68,8%	8,8%	10,0%	7,5%	5,0%	84,6%	81,3%	5,2%	3,2%	3,4%	2,9%	2,4%	4,5%	4,4%	94,6%	65,1%	8,3%
6	81,3%	7,5%	11,3%	0,0%	0,0%	89,0%	91,3%	4,4%	2,9%	3,5%	0,0%	0,0%	3,7%	3,2%	93,3%	75,8%	7,0%
7	88,8%	2,5%	2,5%	2,5%	3,8%	95,9%	92,5%	3,5%	1,7%	1,7%	1,7%	2,1%	2,3%	2,9%	90,1%	79,9%	2,3%
8	95,0%	1,3%	1,3%	1,3%	1,3%	97,4%	97,5%	2,4%	1,2%	1,2%	1,2%	1,2%	1,8%	1,7%	90,8%	86,3%	1,1%
9	90,0%	5,0%	2,5%	2,5%	0,0%	98,6%	91,3%	3,4%	2,4%	1,7%	1,7%	0,0%	1,4%	3,2%	91,7%	82,5%	4,6%
10	91,3%	2,5%	2,5%	3,8%	0,0%	96,1%	95,0%	3,2%	1,7%	1,7%	2,1%	0,0%	2,2%	2,4%	93,4%	85,3%	2,3%
11	87,5%	1,3%	3,8%	7,5%	0,0%	89,7%	97,5%	3,7%	1,2%	2,1%	2,9%	0,0%	3,4%	1,7%	93,3%	81,6%	1,2%
12	83,8%	5,0%	2,5%	1,3%	7,5%	93,1%	90,0%	4,1%	2,4%	1,7%	1,2%	2,9%	3,0%	3,4%	94,6%	79,3%	4,7%
13	85,0%	6,3%	1,3%	3,8%	3,8%	93,2%	91,3%	4,0%	2,7%	1,2%	2,1%	2,1%	3,0%	3,2%	93,6%	79,5%	5,8%
14	88,8%	2,5%	0,0%	3,8%	5,0%	97,3%	91,3%	3,5%	1,7%	0,0%	2,1%	2,4%	1,9%	3,2%	89,0%	79,0%	2,2%
15	91,3%	3,8%	2,5%	1,3%	1,3%	97,3%	93,8%	3,2%	2,1%	1,7%	1,2%	1,2%	1,9%	2,7%	90,5%	82,6%	3,4%
16	85,0%	5,0%	3,8%	2,5%	3,8%	94,4%	90,0%	4,0%	2,4%	2,1%	1,7%	2,1%	2,7%	3,4%	85,9%	73,0%	4,3%
17	85,0%	3,8%	7,5%	0,0%	3,8%	93,2%	91,3%	4,0%	2,1%	2,9%	0,0%	2,1%	3,0%	3,2%	82,3%	69,9%	3,1%
18	73,8%	8,8%	13,8%	2,5%	1,3%	85,5%	86,3%	4,9%	3,2%	3,9%	1,7%	1,2%	4,2%	3,9%	87,9%	64,8%	7,7%
19	65,0%	17,5%	11,3%	2,5%	3,8%	82,5%	78,8%	5,3%	4,2%	3,5%	1,7%	2,1%	4,8%	4,6%	93,9%	61,0%	16,4%
20	65,0%	12,5%	21,3%	0,0%	1,3%	86,7%	75,0%	5,3%	3,7%	4,6%	0,0%	1,2%	4,4%	4,8%	94,2%	61,2%	11,8%
21	71,3%	13,8%	10,0%	3,8%	1,3%	91,9%	77,5%	5,1%	3,9%	3,4%	2,1%	1,2%	3,5%	4,7%	95,0%	67,7%	13,1%
22	66,3%	16,3%	12,5%	1,3%	3,8%	86,9%	76,3%	5,3%	4,1%	3,7%	1,2%	2,1%	4,3%	4,8%	93,7%	62,1%	15,2%
23	61,3%	13,8%	7,5%	13,8%	3,8%	75,4%	81,3%	5,4%	3,9%	2,9%	3,9%	2,1%	5,3%	4,4%	93,0%	57,0%	12,8%

Figura 54 Tabela com os índices e erros associados para SJS 18/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	82,4%	11,8%	0,0%	5,9%	0,0%	82,4%	100,0%	9,2%	7,8%	0,0%	5,7%	0,0%	9,2%	0,0%	4,0%	3,3%	0,5%
1	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
2	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
3	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
5	66,7%	0,0%	33,3%	0,0%	0,0%	66,7%	100,0%	15,7%	0,0%	15,7%	0,0%	0,0%	15,7%	0,0%	2,2%	1,5%	0,0%
6	92,5%	7,5%	0,0%	0,0%	0,0%	92,5%	100,0%	4,2%	4,2%	0,0%	0,0%	0,0%	4,2%	0,0%	31,1%	28,8%	2,3%
7	85,0%	7,5%	7,5%	0,0%	0,0%	85,0%	100,0%	5,6%	4,2%	4,2%	0,0%	0,0%	5,6%	0,0%	35,3%	30,0%	2,6%
8	87,5%	12,5%	0,0%	0,0%	0,0%	87,5%	100,0%	5,2%	5,2%	0,0%	0,0%	0,0%	5,2%	0,0%	33,2%	29,1%	4,2%
9	97,5%	0,0%	2,5%	0,0%	0,0%	97,5%	100,0%	2,5%	0,0%	2,5%	0,0%	0,0%	2,5%	0,0%	32,2%	31,4%	0,0%
10	87,5%	12,5%	0,0%	0,0%	0,0%	87,5%	100,0%	5,2%	5,2%	0,0%	0,0%	0,0%	5,2%	0,0%	30,9%	27,0%	3,9%
11	82,5%	15,0%	2,5%	0,0%	0,0%	82,5%	100,0%	6,0%	5,6%	2,5%	0,0%	0,0%	6,0%	0,0%	33,7%	27,8%	5,0%
12	92,5%	5,0%	2,5%	0,0%	0,0%	92,5%	100,0%	4,2%	3,4%	2,5%	0,0%	0,0%	4,2%	0,0%	34,3%	31,7%	1,7%
13	85,0%	12,5%	2,5%	0,0%	0,0%	85,0%	100,0%	5,6%	5,2%	2,5%	0,0%	0,0%	5,6%	0,0%	31,5%	26,8%	3,9%
14	80,0%	17,5%	2,5%	0,0%	0,0%	80,0%	100,0%	6,3%	6,0%	2,5%	0,0%	0,0%	6,3%	0,0%	30,4%	24,3%	5,3%
15	85,0%	10,0%	5,0%	0,0%	0,0%	85,0%	100,0%	5,6%	4,7%	3,4%	0,0%	0,0%	5,6%	0,0%	32,8%	27,9%	3,3%
16	95,0%	5,0%	0,0%	0,0%	0,0%	95,0%	100,0%	3,4%	3,4%	0,0%	0,0%	0,0%	3,4%	0,0%	30,9%	29,3%	1,5%
17	90,0%	10,0%	0,0%	0,0%	0,0%	90,0%	100,0%	4,7%	4,7%	0,0%	0,0%	0,0%	4,7%	0,0%	34,2%	30,8%	3,4%
18	77,5%	10,0%	12,5%	0,0%	0,0%	77,5%	100,0%	6,6%	4,7%	5,2%	0,0%	0,0%	6,6%	0,0%	11,3%	8,7%	1,1%
19	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
20	0,0%	0,0%	0,0%	100,0%	0,0%	0,0%	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,3%	0,0%	0,0%
21	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
22	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
23	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Figura 55 Tabela com os índices e erros associados para BMA 17/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
1	0,0%	0,0%	0,0%	100,0%	0,0%	0,0%	50,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	35,4%	0,0%	0,0%	0,0%
2	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,8%	0,0%	0,0%
3	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
5	86,7%	13,3%	0,0%	0,0%	0,0%	86,7%	100,0%	8,8%	8,8%	0,0%	0,0%	0,0%	8,8%	0,0%	4,1%	3,6%	0,5%
6	85,0%	10,0%	5,0%	0,0%	0,0%	85,0%	100,0%	5,6%	4,7%	3,4%	0,0%	0,0%	5,6%	0,0%	29,2%	24,8%	2,9%
7	90,0%	7,5%	2,5%	0,0%	0,0%	90,0%	100,0%	4,7%	4,2%	2,5%	0,0%	0,0%	4,7%	0,0%	35,0%	31,5%	2,6%
8	97,5%	2,5%	0,0%	0,0%	0,0%	97,5%	100,0%	2,5%	2,5%	0,0%	0,0%	0,0%	2,5%	0,0%	34,7%	33,8%	0,9%
9	92,5%	7,5%	0,0%	0,0%	0,0%	92,5%	100,0%	4,2%	4,2%	0,0%	0,0%	0,0%	4,2%	0,0%	32,1%	29,7%	2,4%
10	87,5%	12,5%	0,0%	0,0%	0,0%	87,5%	100,0%	5,2%	5,2%	0,0%	0,0%	0,0%	5,2%	0,0%	33,6%	29,4%	4,2%
11	87,5%	7,5%	5,0%	0,0%	0,0%	87,5%	100,0%	5,2%	4,2%	3,4%	0,0%	0,0%	5,2%	0,0%	33,9%	29,6%	2,5%
12	82,5%	12,5%	5,0%	0,0%	0,0%	82,5%	100,0%	6,0%	5,2%	3,4%	0,0%	0,0%	6,0%	0,0%	32,0%	26,4%	4,0%
13	75,0%	10,0%	15,0%	0,0%	0,0%	75,0%	100,0%	6,8%	4,7%	5,6%	0,0%	0,0%	6,8%	0,0%	32,9%	24,7%	3,3%
14	88,8%	10,0%	1,3%	0,0%	0,0%	88,8%	100,0%	3,5%	3,4%	1,2%	0,0%	0,0%	3,5%	0,0%	29,6%	26,2%	3,0%
15	88,8%	11,3%	0,0%	0,0%	0,0%	88,8%	100,0%	3,5%	3,5%	0,0%	0,0%	0,0%	3,5%	0,0%	33,9%	30,1%	3,8%
16	92,5%	7,5%	0,0%	0,0%	0,0%	92,5%	100,0%	2,9%	2,9%	0,0%	0,0%	0,0%	2,9%	0,0%	33,2%	30,7%	2,5%
17	90,0%	7,5%	2,5%	0,0%	0,0%	90,0%	100,0%	3,4%	2,9%	1,7%	0,0%	0,0%	3,4%	0,0%	33,7%	30,4%	2,5%
18	77,5%	12,5%	8,8%	1,3%	0,0%	77,5%	100,0%	4,7%	3,7%	3,2%	1,2%	0,0%	4,7%	0,0%	11,6%	9,0%	1,4%
19	80,0%	13,8%	3,8%	2,5%	0,0%	80,0%	100,0%	4,5%	3,9%	2,1%	1,7%	0,0%	4,5%	0,0%	9,2%	7,4%	1,3%
20	80,0%	11,3%	7,5%	1,3%	0,0%	80,0%	100,0%	4,5%	3,5%	2,9%	1,2%	0,0%	4,5%	0,0%	7,7%	6,1%	0,9%
21	82,5%	8,8%	7,5%	1,3%	0,0%	82,5%	100,0%	4,2%	3,2%	2,9%	1,2%	0,0%	4,2%	0,0%	6,8%	5,6%	0,6%
22	63,9%	25,0%	9,7%	1,4%	0,0%	63,9%	100,0%	5,7%	5,1%	3,5%	1,4%	0,0%	5,7%	0,0%	5,0%	3,2%	1,3%
23	69,9%	19,2%	9,6%	1,4%	0,0%	69,9%	100,0%	5,4%	4,6%	3,4%	1,4%	0,0%	5,4%	0,0%	6,3%	4,4%	1,2%

Figura 56 Tabela com os índices e erros associados para BMA 18/9/2012.



hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
1	19,4%	11,9%	16,4%	52,2%	0,0%	92,9%	20,9%	4,8%	4,0%	4,5%	6,1%	0,0%	6,9%	5,0%	13,4%	2,6%	1,6%
2	14,7%	2,9%	26,5%	55,9%	0,0%	83,3%	17,6%	6,1%	2,9%	7,6%	8,5%	0,0%	15,2%	6,5%	11,6%	1,7%	0,3%
3	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
4	16,1%	5,4%	12,5%	66,1%	0,0%	90,0%	17,9%	4,9%	3,0%	4,4%	6,3%	0,0%	9,5%	5,1%	60,0%	9,6%	3,2%
5	14,7%	8,0%	6,7%	70,7%	0,0%	84,6%	17,3%	4,1%	3,1%	2,9%	5,3%	0,0%	10,0%	4,4%	99,7%	14,6%	8,0%
6	27,5%	16,3%	10,0%	46,3%	0,0%	75,9%	36,3%	5,0%	4,1%	3,4%	5,6%	0,0%	7,9%	5,4%	99,8%	27,4%	16,2%
7	31,0%	23,0%	10,0%	36,0%	0,0%	93,9%	33,0%	4,6%	4,2%	3,0%	4,8%	0,0%	4,2%	4,7%	97,9%	30,3%	22,5%
8	27,3%	22,3%	10,7%	39,7%	0,0%	73,3%	37,2%	4,0%	3,8%	2,8%	4,4%	0,0%	6,6%	4,4%	100,0%	27,3%	22,3%
9	32,5%	28,3%	10,8%	28,3%	0,0%	86,7%	37,5%	4,3%	4,1%	2,8%	4,1%	0,0%	5,1%	4,4%	100,0%	32,5%	28,3%
10	28,9%	21,5%	19,8%	29,8%	0,0%	100,0%	28,9%	4,1%	3,7%	3,6%	4,2%	0,0%	0,0%	4,1%	99,9%	28,9%	21,5%
11	13,3%	30,0%	22,5%	34,2%	0,0%	72,7%	18,3%	3,1%	4,2%	3,8%	4,3%	0,0%	9,5%	3,5%	99,9%	13,3%	30,0%
12	30,0%	28,3%	11,7%	30,0%	0,0%	85,7%	35,0%	4,2%	4,1%	2,9%	4,2%	0,0%	5,4%	4,4%	99,8%	29,9%	28,3%
13	24,8%	26,4%	19,8%	28,9%	0,0%	85,7%	28,9%	3,9%	4,0%	3,6%	4,1%	0,0%	5,9%	4,1%	99,9%	24,8%	26,4%
14	17,4%	38,0%	21,5%	23,1%	0,0%	77,8%	22,3%	3,4%	4,4%	3,7%	3,8%	0,0%	8,0%	3,8%	99,8%	17,3%	37,9%
15	35,8%	35,8%	21,7%	6,7%	0,0%	89,6%	40,0%	4,4%	4,4%	3,8%	2,3%	0,0%	4,4%	4,5%	99,8%	35,8%	35,8%
16	39,2%	31,7%	17,5%	11,7%	0,0%	94,0%	41,7%	4,5%	4,2%	3,5%	2,9%	0,0%	3,4%	4,5%	99,9%	39,1%	31,6%
17	38,3%	26,7%	14,2%	20,8%	0,0%	90,2%	42,5%	4,4%	4,0%	3,2%	3,7%	0,0%	4,2%	4,5%	99,9%	38,3%	26,6%
18	19,8%	4,1%	24,0%	52,1%	0,0%	88,9%	22,3%	3,6%	1,8%	3,9%	4,5%	0,0%	6,0%	3,8%	99,8%	19,8%	4,1%
19	18,3%	8,3%	9,2%	64,2%	0,0%	91,7%	20,0%	3,5%	2,5%	2,6%	4,4%	0,0%	5,6%	3,7%	99,8%	18,3%	8,3%
20	18,0%	7,4%	8,2%	66,4%	0,0%	91,7%	19,7%	3,5%	2,4%	2,5%	4,3%	0,0%	5,6%	3,6%	99,8%	18,0%	7,4%
21	13,3%	8,3%	5,0%	73,3%	0,0%	84,2%	15,8%	3,1%	2,5%	2,0%	4,0%	0,0%	8,4%	3,3%	99,9%	13,3%	8,3%
22	19,7%	11,1%	4,3%	65,0%	0,0%	82,1%	23,9%	3,7%	2,9%	1,9%	4,4%	0,0%	7,2%	3,9%	99,8%	19,6%	11,1%
23	17,5%	8,3%	5,0%	69,2%	0,0%	91,3%	19,2%	3,5%	2,5%	2,0%	4,2%	0,0%	5,9%	3,6%	99,7%	17,4%	8,3%

Figura 57 Tabela com os índices e erros associados para SJM 17/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	20,0%	10,5%	2,9%	66,7%	0,0%	87,5%	22,9%	3,9%	3,0%	1,6%	4,6%	0,0%	6,8%	4,1%	98,2%	19,6%	10,3%
1	11,7%	6,5%	6,5%	75,3%	0,0%	81,8%	14,3%	3,7%	2,8%	2,8%	4,9%	0,0%	11,6%	4,0%	97,5%	11,4%	6,3%
2	17,2%	7,8%	3,1%	71,9%	0,0%	91,7%	18,8%	4,7%	3,4%	2,2%	5,6%	0,0%	8,0%	4,9%	100,0%	17,2%	7,8%
3	13,3%	8,3%	6,7%	71,7%	0,0%	88,9%	15,0%	4,4%	3,6%	3,2%	5,8%	0,0%	10,5%	4,6%	98,1%	13,1%	8,2%
4	9,1%	12,1%	9,1%	69,7%	0,0%	60,0%	15,2%	3,5%	4,0%	3,5%	5,7%	0,0%	15,5%	4,4%	96,9%	8,8%	11,7%
5	19,4%	8,3%	8,3%	63,9%	0,0%	87,5%	22,2%	4,7%	3,3%	3,3%	5,7%	0,0%	8,3%	4,9%	98,5%	19,1%	8,2%
6	21,3%	17,5%	13,8%	47,5%	0,0%	89,5%	23,8%	4,6%	4,2%	3,9%	5,6%	0,0%	7,0%	4,8%	99,7%	21,2%	17,4%
7	32,0%	29,0%	4,0%	35,0%	0,0%	94,1%	34,0%	4,7%	4,5%	2,0%	4,8%	0,0%	4,0%	4,7%	99,6%	31,9%	28,9%
8	36,7%	25,0%	10,8%	27,5%	0,0%	93,6%	39,2%	4,4%	4,0%	2,8%	4,1%	0,0%	3,6%	4,5%	99,7%	36,6%	24,9%
9	27,8%	32,3%	16,5%	23,3%	0,0%	88,1%	31,6%	3,9%	4,1%	3,2%	3,7%	0,0%	5,0%	4,0%	99,2%	27,6%	32,1%
10	26,0%	31,3%	21,4%	20,6%	0,8%	81,0%	32,1%	3,8%	4,1%	3,6%	3,5%	0,8%	6,1%	4,1%	99,9%	25,9%	31,3%
11	21,7%	30,0%	19,2%	29,2%	0,0%	89,7%	24,2%	3,8%	4,2%	3,6%	4,1%	0,0%	5,7%	3,9%	100,0%	21,7%	30,0%
12	28,3%	29,2%	25,0%	17,5%	0,0%	85,0%	33,3%	4,1%	4,1%	4,0%	3,5%	0,0%	5,6%	4,3%	99,7%	28,3%	29,1%
13	25,8%	32,5%	9,2%	32,5%	0,0%	86,1%	30,0%	4,0%	4,3%	2,6%	4,3%	0,0%	5,8%	4,2%	99,4%	25,7%	32,3%
14	27,5%	35,8%	11,7%	25,0%	0,0%	91,7%	30,0%	4,1%	4,4%	2,9%	4,0%	0,0%	4,6%	4,2%	97,9%	26,9%	35,1%
15	25,4%	24,6%	18,0%	32,0%	0,0%	86,1%	29,5%	3,9%	3,9%	3,5%	4,2%	0,0%	5,8%	4,1%	94,0%	23,9%	23,1%
16	37,7%	23,0%	13,1%	26,2%	0,0%	90,2%	41,8%	4,4%	3,8%	3,1%	4,0%	0,0%	4,2%	4,5%	91,5%	34,5%	21,0%
17	46,7%	28,7%	8,2%	16,4%	0,0%	91,9%	50,8%	4,5%	4,1%	2,5%	3,4%	0,0%	3,5%	4,5%	94,9%	44,3%	27,2%
18	14,3%	5,6%	7,1%	73,0%	0,0%	85,7%	16,7%	3,1%	2,0%	2,3%	4,0%	0,0%	7,6%	3,3%	98,6%	14,1%	5,5%
19	14,2%	9,2%	5,8%	70,8%	0,0%	81,0%	17,5%	3,2%	2,6%	2,1%	4,1%	0,0%	8,6%	3,5%	99,6%	14,1%	9,1%
20	17,4%	6,6%	6,6%	69,4%	0,0%	87,5%	19,8%	3,4%	2,3%	2,3%	4,2%	0,0%	6,8%	3,6%	99,2%	17,2%	6,6%
21	27,3%	9,1%	5,0%	58,7%	0,0%	91,7%	29,8%	4,0%	2,6%	2,0%	4,5%	0,0%	4,6%	4,2%	99,5%	27,1%	9,0%
22	14,5%	8,5%	6,0%	70,9%	0,0%	89,5%	16,2%	3,3%	2,6%	2,2%	4,2%	0,0%	7,0%	3,4%	99,4%	14,4%	8,5%
23	19,1%	7,0%	0,0%	73,9%	0,0%	88,0%	21,7%	3,7%	2,4%	0,0%	4,1%	0,0%	6,5%	3,8%	98,5%	18,8%	6,8%

Figura 58 Tabela com os índices e erros associados para SJM 18/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	69,2%	7,7%	0,0%	23,1%	0,0%	69,2%	100,0%	5,2%	3,0%	0,0%	4,8%	0,0%	5,2%	0,0%	31,4%	21,8%	2,4%
1	64,5%	9,2%	2,6%	23,7%	0,0%	64,5%	100,0%	5,5%	3,3%	1,8%	4,9%	0,0%	5,5%	0,0%	39,4%	25,4%	3,6%
2	46,2%	3,1%	4,6%	46,2%	0,0%	46,2%	100,0%	6,2%	2,1%	2,6%	6,2%	0,0%	6,2%	0,0%	40,7%	18,8%	1,3%
3	46,5%	8,5%	5,6%	39,4%	0,0%	46,5%	100,0%	5,9%	3,3%	2,7%	5,8%	0,0%	5,9%	0,0%	51,5%	23,9%	4,4%
4	42,1%	6,6%	2,6%	48,7%	0,0%	42,1%	100,0%	5,7%	2,8%	1,8%	5,7%	0,0%	5,7%	0,0%	39,5%	16,6%	2,6%
5	60,3%	5,1%	3,8%	30,8%	0,0%	60,3%	100,0%	5,5%	2,5%	2,2%	5,2%	0,0%	5,5%	0,0%	28,6%	17,2%	1,5%
6	60,3%	3,8%	2,6%	33,3%	0,0%	60,3%	100,0%	5,5%	2,2%	1,8%	5,3%	0,0%	5,5%	0,0%	29,6%	17,9%	1,1%
7	74,4%	1,3%	1,3%	23,1%	0,0%	74,4%	100,0%	4,9%	1,3%	1,3%	4,8%	0,0%	4,9%	0,0%	25,5%	19,0%	0,3%
8	70,5%	1,3%	1,3%	26,9%	0,0%	70,5%	100,0%	5,2%	1,3%	1,3%	5,0%	0,0%	5,2%	0,0%	27,8%	19,6%	0,4%
9	73,1%	6,4%	7,7%	12,8%	0,0%	73,1%	100,0%	5,0%	2,8%	3,0%	3,8%	0,0%	5,0%	0,0%	30,9%	22,6%	2,0%
10	59,0%	9,0%	9,0%	23,1%	0,0%	59,0%	100,0%	5,6%	3,2%	3,2%	4,8%	0,0%	5,6%	0,0%	29,5%	17,4%	2,6%
11	67,9%	3,8%	11,5%	16,7%	0,0%	67,9%	100,0%	5,3%	2,2%	3,6%	4,2%	0,0%	5,3%	0,0%	32,0%	21,7%	1,2%
12	70,5%	6,4%	9,0%	14,1%	0,0%	70,5%	100,0%	5,2%	2,8%	3,2%	3,9%	0,0%	5,2%	0,0%	31,2%	22,0%	2,0%
13	74,4%	2,6%	9,0%	14,1%	0,0%	74,4%	100,0%	4,9%	1,8%	3,2%	3,9%	0,0%	4,9%	0,0%	34,5%	25,7%	0,9%
14	70,5%	12,8%	3,8%	12,8%	0,0%	70,5%	100,0%	5,2%	3,8%	2,2%	3,8%	0,0%	5,2%	0,0%	53,6%	37,8%	6,9%
15	75,6%	6,4%	5,1%	12,8%	0,0%	75,6%	100,0%	4,9%	2,8%	2,5%	3,8%	0,0%	4,9%	0,0%	67,4%	51,0%	4,3%
16	87,2%	5,1%	2,6%	5,1%	0,0%	87,2%	100,0%	3,8%	2,5%	1,8%	2,5%	0,0%	3,8%	0,0%	69,3%	60,4%	3,6%
17	79,5%	10,3%	6,4%	3,8%	0,0%	79,5%	100,0%	4,6%	3,4%	2,8%	2,2%	0,0%	4,6%	0,0%	41,0%	32,6%	4,2%
18	65,4%	6,4%	9,0%	19,2%	0,0%	65,4%	100,0%	5,4%	2,8%	3,2%	4,5%	0,0%	5,4%	0,0%	28,6%	18,7%	1,8%
19	61,5%	9,0%	10,3%	19,2%	0,0%	61,5%	100,0%	5,5%	3,2%	3,4%	4,5%	0,0%	5,5%	0,0%	29,5%	18,1%	2,6%
20	69,2%	5,1%	7,7%	17,9%	0,0%	69,2%	100,0%	5,2%	2,5%	3,0%	4,3%	0,0%	5,2%	0,0%	31,8%	22,0%	1,6%
21	66,7%	2,6%	12,8%	17,9%	0,0%	66,7%	100,0%	5,3%	1,8%	3,8%	4,3%	0,0%	5,3%	0,0%	32,9%	21,9%	0,8%
22	61,5%	2,6%	3,8%	32,1%	0,0%	61,5%	100,0%	5,5%	1,8%	2,2%	5,3%	0,0%	5,5%	0,0%	32,1%	19,8%	0,8%
23	60,3%	5,1%	6,4%	28,2%	0,0%	60,3%	100,0%	5,5%	2,5%	2,8%	5,1%	0,0%	5,5%	0,0%	39,6%	23,9%	2,0%

Figura 59 Tabela com os índices e erros associados para SJC 17/9/2012.

hora	%fCE	%fCH	%fPI	%fPP	%fT	%fPAL	%fIL	sfCE	sfCH	sfPI	sfPP	sfT	sfPAL	sfIL	1-IRV	%cfCE	%cfCH
0	62,8%	3,8%	2,6%	30,8%	0,0%	62,8%	100,0%	5,5%	2,2%	1,8%	5,2%	0,0%	5,5%	0,0%	39,7%	25,0%	1,5%
1	53,2%	3,9%	1,3%	41,6%	0,0%	53,2%	100,0%	5,7%	2,2%	1,3%	5,6%	0,0%	5,7%	0,0%	36,8%	19,6%	1,4%
2	50,7%	8,5%	1,4%	39,4%	0,0%	50,7%	100,0%	5,9%	3,3%	1,4%	5,8%	0,0%	5,9%	0,0%	43,3%	21,9%	3,7%
3	36,4%	6,1%	4,5%	53,0%	0,0%	36,4%	100,0%	5,9%	2,9%	2,6%	6,1%	0,0%	5,9%	0,0%	52,0%	18,9%	3,2%
4	49,3%	5,5%	4,1%	41,1%	0,0%	49,3%	100,0%	5,9%	2,7%	2,3%	5,8%	0,0%	5,9%	0,0%	43,5%	21,5%	2,4%
5	50,0%	6,4%	10,3%	33,3%	0,0%	50,0%	100,0%	5,7%	2,8%	3,4%	5,3%	0,0%	5,7%	0,0%	32,7%	16,4%	2,1%
6	61,5%	2,6%	9,0%	26,9%	0,0%	61,5%	100,0%	5,5%	1,8%	3,2%	5,0%	0,0%	5,5%	0,0%	30,6%	18,8%	0,8%
7	69,2%	3,8%	7,7%	19,2%	0,0%	69,2%	100,0%	5,2%	2,2%	3,0%	4,5%	0,0%	5,2%	0,0%	26,7%	18,5%	1,0%
8	66,7%	7,7%	7,7%	17,9%	0,0%	66,7%	100,0%	5,3%	3,0%	3,0%	4,3%	0,0%	5,3%	0,0%	27,9%	18,6%	2,1%
9	66,7%	6,4%	3,8%	23,1%	0,0%	66,7%	100,0%	5,3%	2,8%	2,2%	4,8%	0,0%	5,3%	0,0%	28,6%	19,1%	1,8%
10	69,2%	5,1%	9,0%	16,7%	0,0%	69,2%	100,0%	5,2%	2,5%	3,2%	4,2%	0,0%	5,2%	0,0%	29,7%	20,6%	1,5%
11	70,5%	3,8%	5,1%	20,5%	0,0%	70,5%	100,0%	5,2%	2,2%	2,5%	4,6%	0,0%	5,2%	0,0%	32,4%	22,9%	1,2%
12	69,2%	7,7%	6,4%	16,7%	0,0%	70,1%	98,7%	5,2%	3,0%	2,8%	4,2%	0,0%	5,2%	1,3%	30,6%	21,2%	2,4%
13	67,9%	3,8%	7,7%	20,5%	0,0%	67,9%	100,0%	5,3%	2,2%	3,0%	4,6%	0,0%	5,3%	0,0%	34,5%	23,5%	1,3%
14	70,5%	5,1%	12,8%	11,5%	0,0%	70,5%	100,0%	5,2%	2,5%	3,8%	3,6%	0,0%	5,2%	0,0%	52,9%	37,3%	2,7%
15	64,1%	14,1%	5,1%	16,7%	0,0%	64,9%	98,7%	5,4%	3,9%	2,5%	4,2%	0,0%	5,4%	1,3%	60,2%	38,6%	8,5%
16	78,2%	5,1%	6,4%	10,3%	0,0%	78,2%	100,0%	4,7%	2,5%	2,8%	3,4%	0,0%	4,7%	0,0%	59,7%	46,7%	3,1%
17	71,8%	3,8%	7,7%	16,7%	0,0%	71,8%	100,0%	5,1%	2,2%	3,0%	4,2%	0,0%	5,1%	0,0%	35,5%	25,5%	1,4%
18	69,2%	7,7%	1,3%	21,8%	0,0%	69,2%	100,0%	5,2%	3,0%	1,3%	4,7%	0,0%	5,2%	0,0%	28,2%	19,5%	2,2%
19	67,9%	6,4%	7,7%	17,9%	0,0%	67,9%	100,0%	5,3%	2,8%	3,0%	4,3%	0,0%	5,3%	0,0%	31,3%	21,3%	2,0%
20	62,8%	5,1%	7,7%	24,4%	0,0%	62,8%	100,0%	5,5%	2,5%	3,0%	4,9%	0,0%	5,5%	0,0%	33,2%	20,8%	1,7%
21	57,7%	6,4%	5,1%	30,8%	0,0%	57,7%	100,0%	5,6%	2,8%	2,5%	5,2%	0,0%	5,6%	0,0%	31,1%	17,9%	2,0%
22	66,7%	1,3%	3,8%	28,2%	0,0%	66,7%	100,0%	5,3%	1,3%	2,2%	5,1%	0,0%	5,3%	0,0%	30,3%	20,2%	0,4%
23	64,1%	2,6%	2,6%	30,8%	0,0%	64,1%	100,0%	5,4%	1,8%	1,8%	5,2%	0,0%	5,4%	0,0%	38,9%	24,9%	1,0%

Figura 60 Tabela com os índices e erros associados para SJC 18/9/2012.

Para facilitar a visualização dos gráficos, no que segue, optamos por não incluir as barras de erro junto a cada ponto com as notas. Os erros associados a cada período e fração estão já mostrados nas tabelas anteriores. Os gráficos permitem observar a evolução temporal de cada nota conforme o período horário.



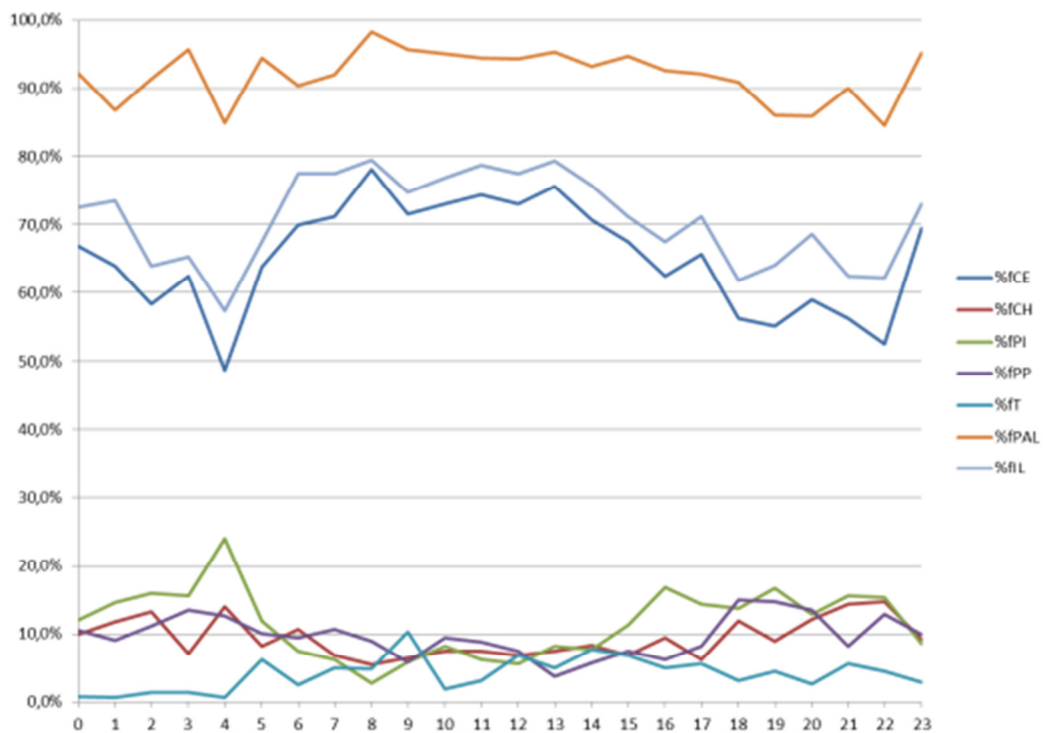


Figura 61 Evolução temporal das diversas notas para GRU em 17/9/2012. Os erros associados não são exibidos para facilitar a leitura.

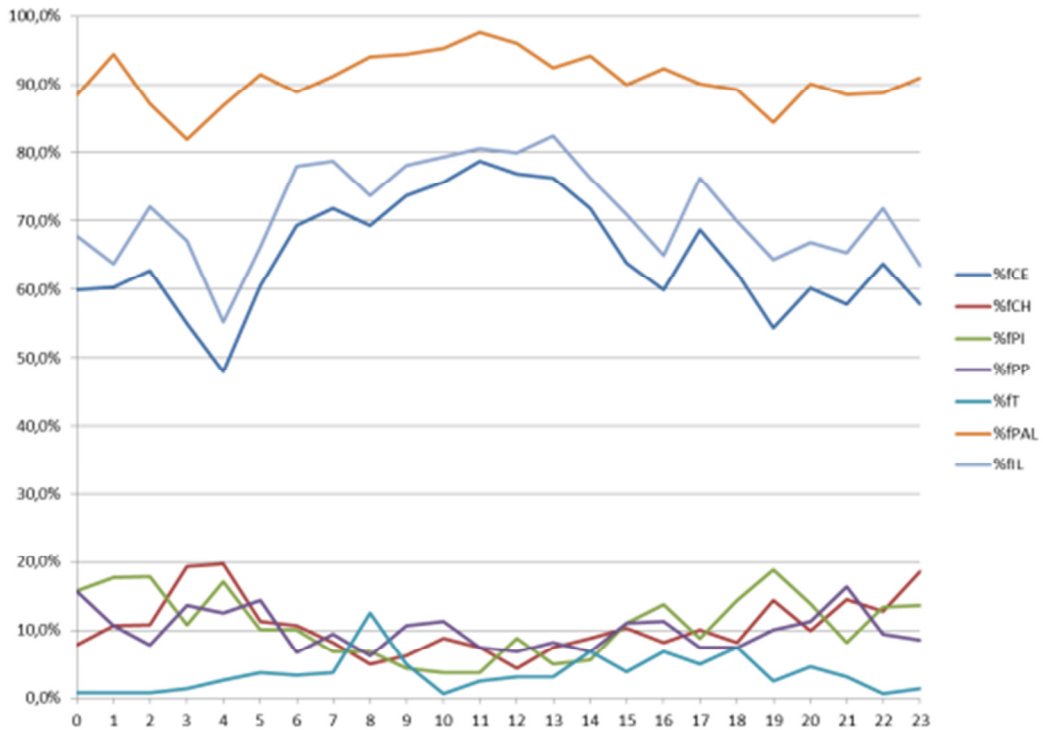


Figura 62 Evolução temporal das diversas notas para GRU em 18/9/2012. Os erros associados não são exibidos para facilitar a leitura.

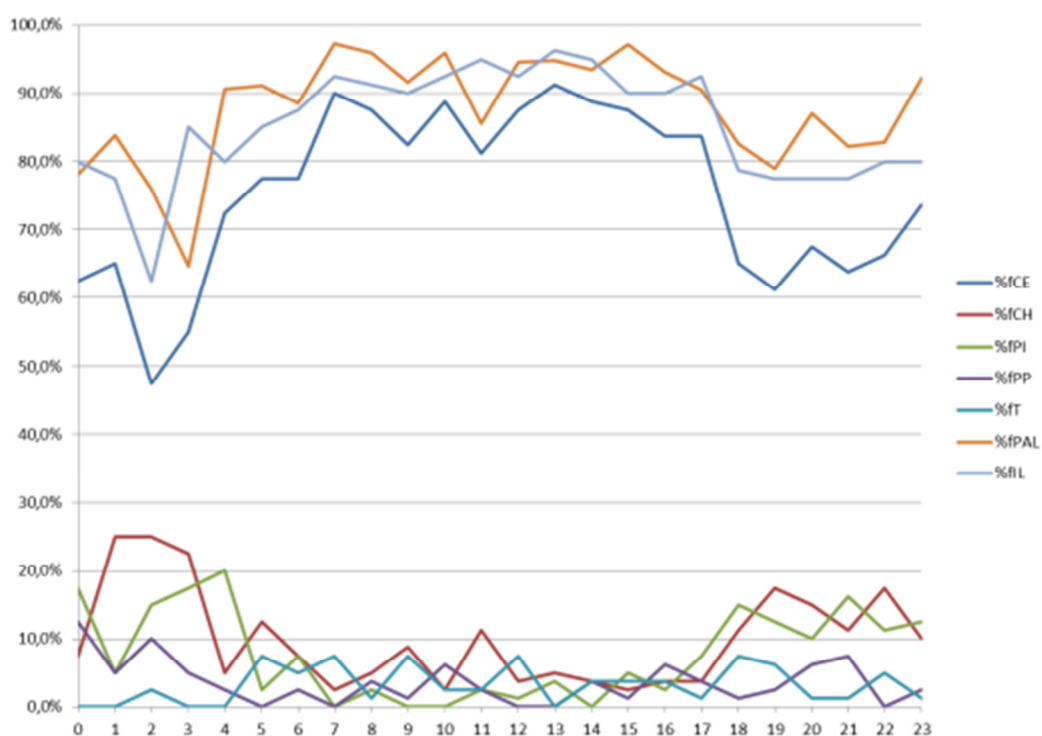


Figura 63 Evolução temporal das diversas notas para SJS em 17/9/2012. Os erros associados não são exibidos para facilitar a leitura.

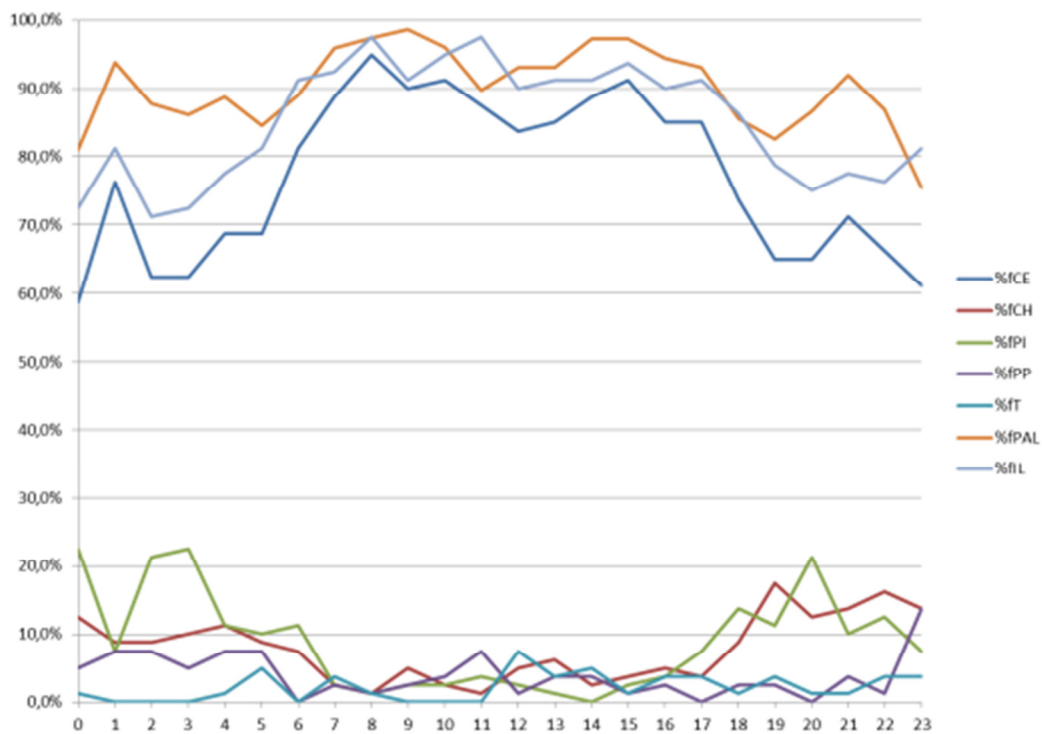


Figura 64 Evolução temporal das diversas notas para SJS em 18/9/2012. Os erros associados não são exibidos para facilitar a leitura.

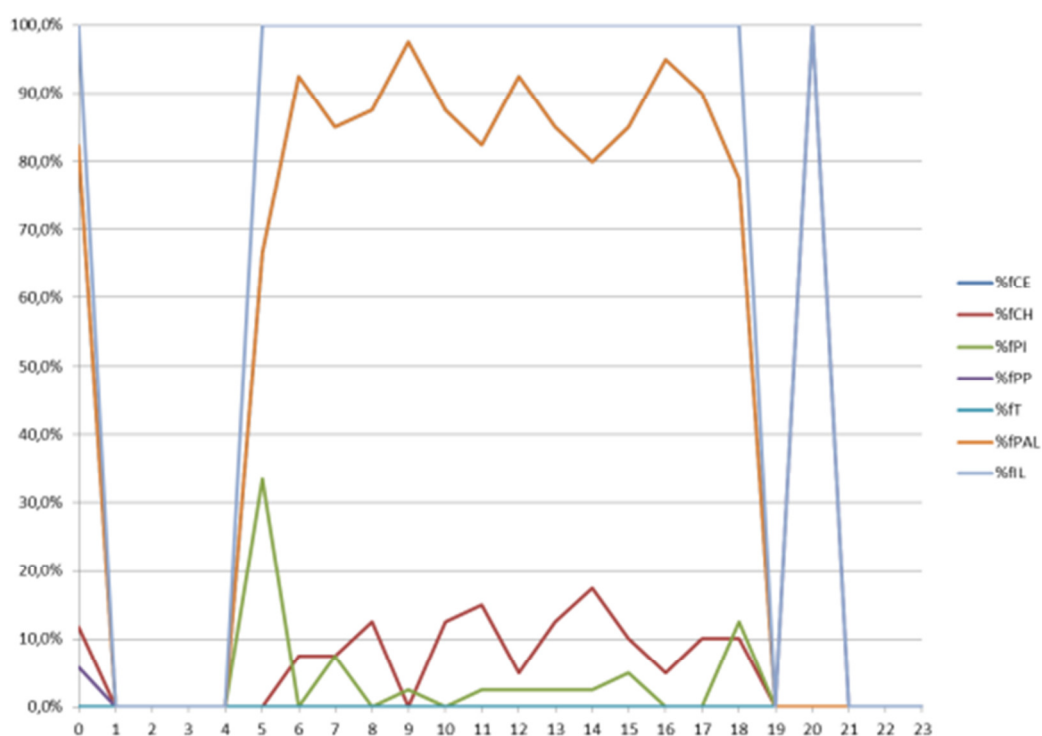


Figura 65 Evolução temporal das diversas notas para BMA em 17/9/2012. Os erros associados não são exibidos para facilitar a leitura.

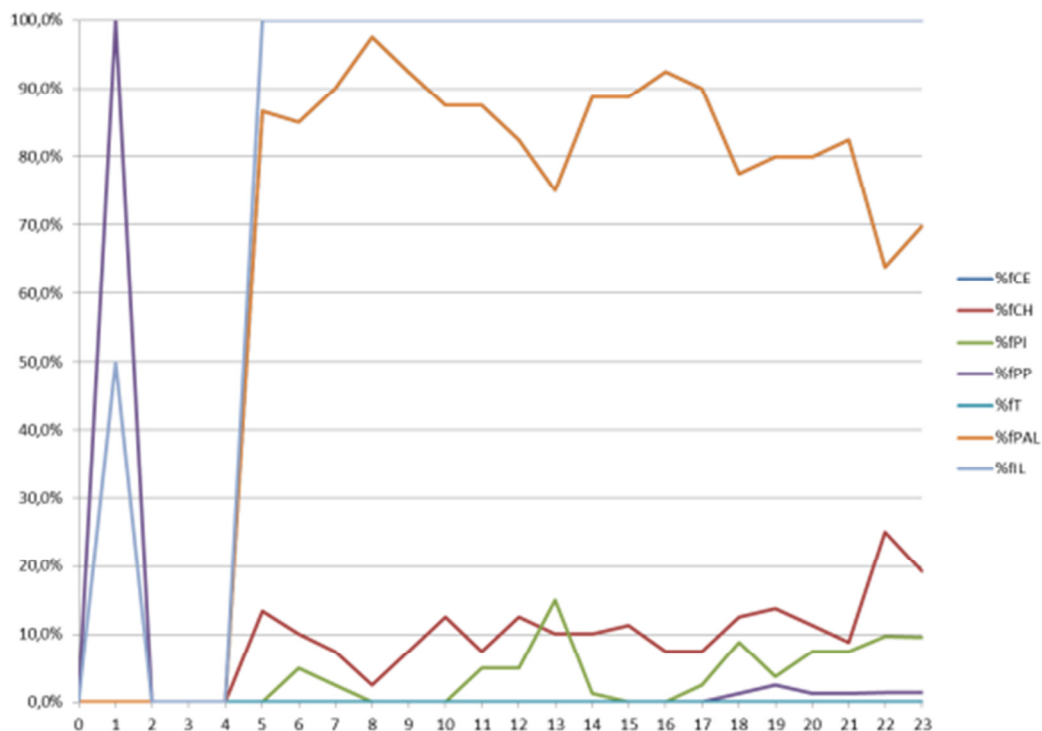


Figura 66 Evolução temporal das diversas notas para BMA em 18/9/2012. Os erros associados não são exibidos para facilitar a leitura.

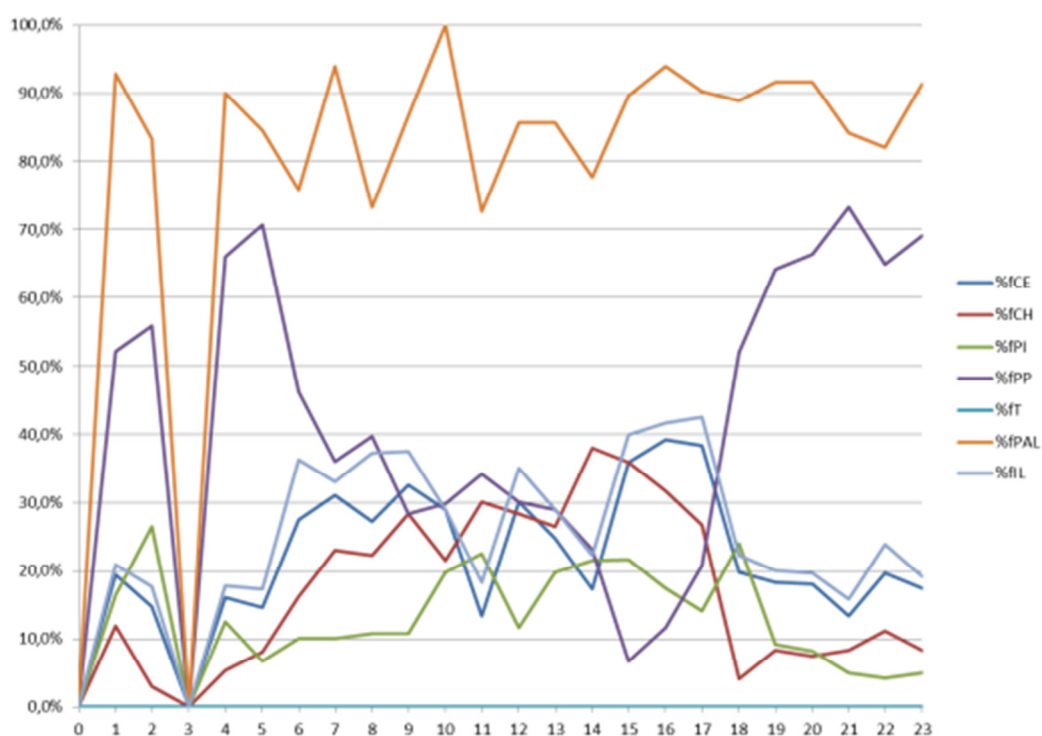


Figura 67 Evolução temporal das diversas notas para SJM em 17/9/2012. Os erros associados não são exibidos para facilitar a leitura.

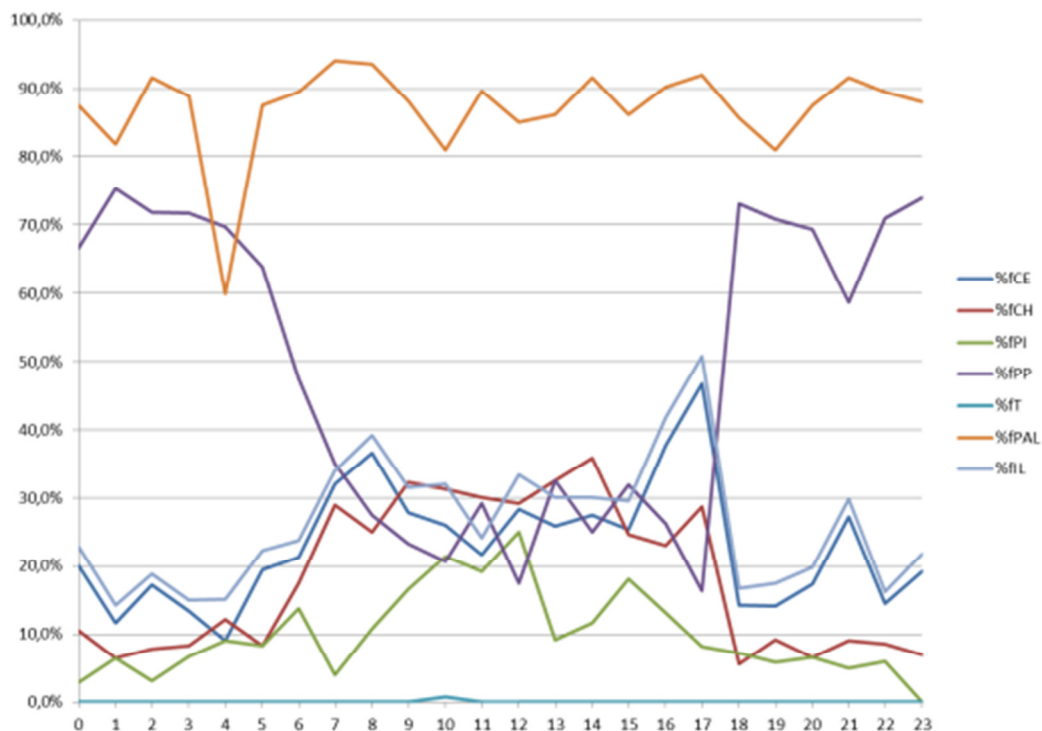


Figura 68 Evolução temporal das diversas notas para SJM em 18/9/2012. Os erros associados não são exibidos para facilitar a leitura.

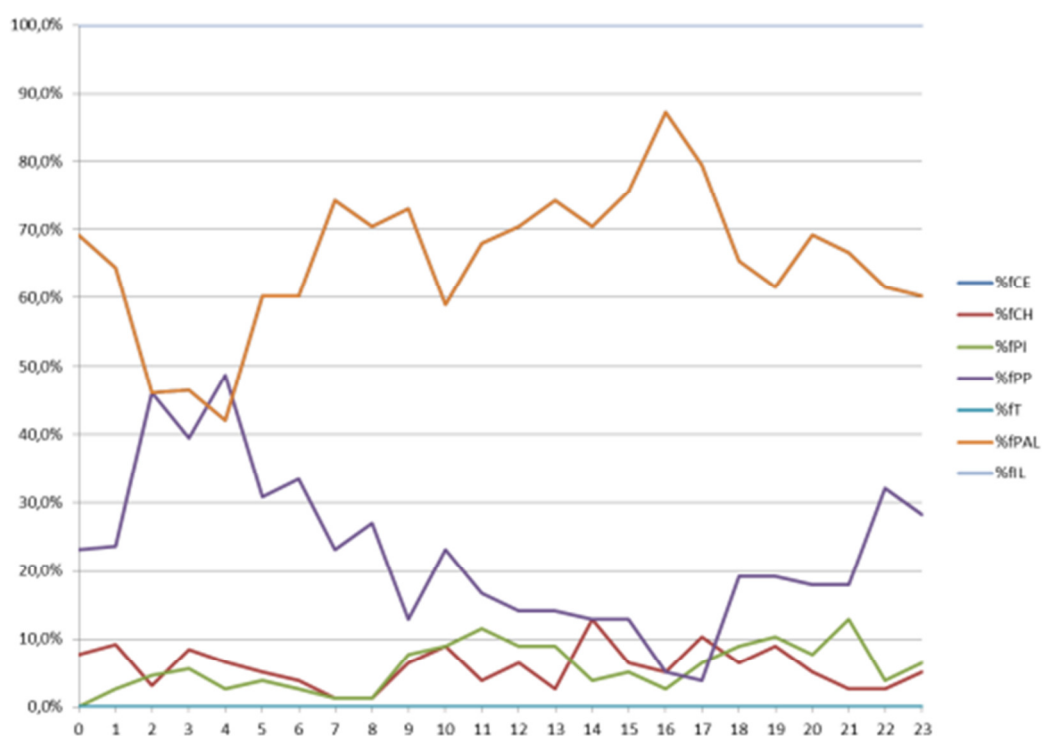


Figura 69 Evolução temporal das diversas notas para SJC em 17/9/2012. Os erros associados não são exibidos para facilitar a leitura.

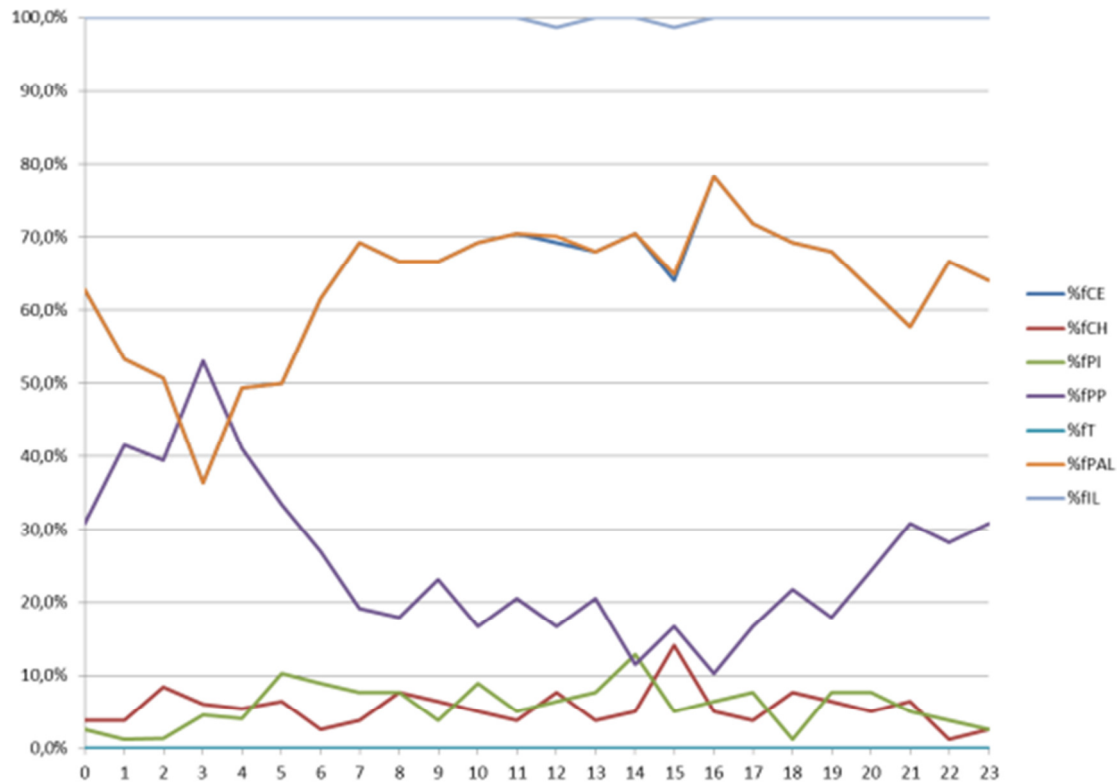


Figura 70 Evolução temporal das diversas notas para SJC em 18/9/2012. Os erros associados não são exibidos para facilitar a leitura.

O índice principal de acerto médio, corrigido para o fluxo – aqui chamado de %cfCE para todos os sítios é a taxa de acerto *considerando toda perda de tráfego*.

	GRU 17/9		GRU 18/9	
Índice	TOTAL	%ERR	TOTAL	%ERR
%fCE	65,29%	0,78%	65,19%	0,78%
%fCH	9,36%	0,48%	10,44%	0,50%
%fPI	11,52%	0,52%	10,68%	0,51%
%fPP	9,65%	0,49%	10,01%	0,49%
%fT	4,18%	0,33%	3,68%	0,31%
%fPAL	92,12%	0,53%	91,08%	0,55%
%fIL	70,87%	0,75%	71,57%	0,74%
%IrV	4,27%		3,91%	
%cfCE	62,50%		62,64%	
%cfCH	8,96%		10,03%	

	SJS 17/9		SJS 18/9	
Índice	TOTAL	%ERR	TOTAL	%ERR
%fCE	76,69%	1,06%	77,19%	0,96%
%fCH	9,31%	0,73%	7,86%	0,61%
%fPI	7,25%	0,65%	8,85%	0,65%
%fPP	3,38%	0,45%	3,91%	0,44%
%fT	3,38%	0,45%	2,19%	0,33%
%fPAL	89,11%	0,84%	90,64%	0,72%
%fIL	86,06%	0,87%	85,16%	0,81%
%IrV	19,30%		8,96%	
%cfCE	61,89%		70,27%	
%cfCH	7,52%		7,16%	

	BMA 17/9		BMA 18/9	
Índice	TOTAL	%ERR	TOTAL	%ERR
%fCE	86,68%	1,45%	83,16%	1,12%
%fCH	9,49%	1,25%	11,41%	0,95%
%fPI	3,28%	0,76%	4,63%	0,63%
%fPP	0,55%	0,32%	0,80%	0,27%
%fT	0,00%	0,00%	0,00%	0,00%
%fPAL	86,68%	1,45%	83,23%	1,12%
%fIL	100,00%	0,00%	99,91%	0,09%
%IrV	77,10%		75,60%	
%cfCE	19,85%		20,29%	
%cfCH	2,17%		2,78%	

	SJM 17/9		SJM 18/9	
Índice	TOTAL	%ERR	TOTAL	%ERR
%fCE	24,19%	0,89%	24,05%	0,84%
%fCH	19,65%	0,82%	19,35%	0,78%
%fPI	13,78%	0,71%	10,37%	0,60%
%fPP	42,38%	1,02%	46,19%	0,98%
%fT	0,00%	0,00%	0,04%	0,04%
%fPAL	86,92%	1,32%	88,18%	1,22%
%fIL	27,83%	0,93%	27,27%	0,88%
%IrV	3,05%		1,54%	
%cfCE	23,45%		23,68%	
%cfCH	19,05%		19,05%	

	<b>SJC 17/9</b>		<b>SJC 18/9</b>	
<b>Índice</b>	<b>TOTAL</b>	<b>%ERR</b>	<b>TOTAL</b>	<b>%ERR</b>
<b>%fCE</b>	65,91%	1,10%	63,45%	1,12%
<b>%fCH</b>	5,84%	0,55%	5,52%	0,53%
<b>%fPI</b>	5,79%	0,54%	5,90%	0,55%
<b>%fPP</b>	22,46%	0,97%	25,12%	1,01%
<b>%fT</b>	0,00%	0,00%	0,00%	0,00%
<b>%fPAL</b>	65,91%	1,10%	63,52%	1,12%
<b>%fIL</b>	100,00%	0,00%	99,89%	0,08%
<b>%IrV</b>	63,20%		64,20%	
<b>%cfCE</b>	24,25%		22,72%	
<b>%cfCH</b>	2,15%		1,98%	

Figura 71 Tabelas contendo as notas médias para os diversos sítios nos dias indicados. Destacamos os índices %Irv (perda de fluxo) e a fração corrigida média de acerto que, no final, o percentual total de veículos efetivamente lidos e validados por cada um dos sistemas.

## Estatísticas específicas

### Leitura automática de motos

Para um dos sítios (SJS) a orientação das câmeras foi para a traseira de veículos, o que permitiu investigar o índice de captura automática de placas de motos. Como o número de ocorrência de motos é muito pequeno ( $N=95$ ), optamos por fazer a estatística integrada para os dias 17/9 e 18/9, pelo que encontramos os índices:

- $\%fCE = 26,3\% \pm 4,5\%$ ;
- $\%fCH = 54,7\% \pm 5,1\%$ ;
- $\%fPI = 17,9\% \pm 3,9\%$ ;
- $\%fPP = 1,1\% \pm 1,0\%$ ;
- $\%fPAL = 64,1\% \pm 7,7\%$ ;
- $\%fIL = 41,1\% \pm 5\%$ .

Portanto, a taxa máxima de acerto de motos é da ordem de 26,3%, sendo que a maioria (54,7%) somente pode ser inspecionada por humanos. Observe também que não é possível determinar o índice corrigido segundo %IrV, pois a análise de tráfego não distingue entre tipos de veículos.

### Análise de um período com chuva (5:00 de 21/9/2012).

Um período foi escolhido com incidência de chuva moderada. O final da madrugada de 21/9/2012 apresentou chuva no sítio GRU que foi usado para se determinar os índices de desempenho que foram:

- $\%fCE = 58,0\% \pm 1,5\%$ ;
- $\%fCH = 7,6\% \pm 0,8\%$ ;
- $\%fPI = 14,0\% \pm 1,0\%$ ;



- $\%fPP=16,3\%\pm 1,1\%$ ;
- $\%fT=4,2\%\pm 0,6\%$ ;
- $\%fPAL=86,7\%\pm 1,2\%$ ;
- $\%fIL=66,9\%\pm 1,4\%$ .

A Figura 72 traz a comparação com um período equivalente do dia 17/9 onde os níveis de erro indicam a equivalência dos resultados para todas as notas.

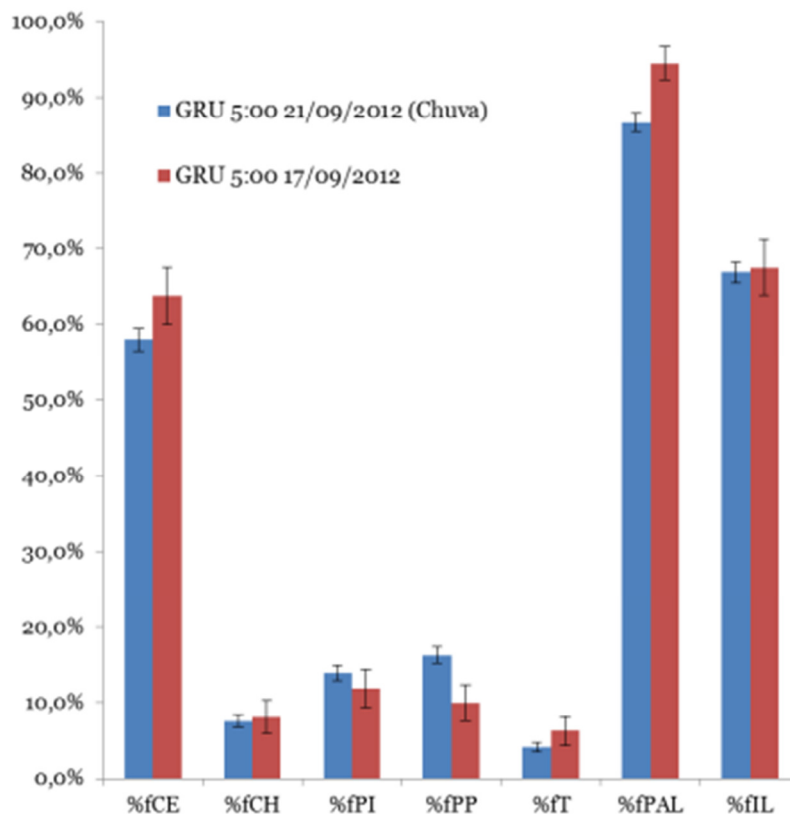


Figura 72 Comparação entre um período normal (vermelho) com o mesmo período em um dia chuvoso (azul). Os dados foram integrados a partir do sítio GRU entre os dias 17/9 e 21/9 (período das 5:00).

Apesar do aumento das frações de perda ( $\%fPI$  e  $\%fPP$ ), não é possível afirmar que a variação observada se deve ao aumento da chuva ou a uma variação normal entre períodos. Assim, em princípio, para a chuva observada no período das 5:00 de 21/9, não é possível afirmar a incidência de qualquer efeito sobre a taxa de acerto do sistema.

## Dimensão estimada de BackOffice

A partir das notas obtidas é possível determinar qual a dimensão de trabalho (em unidades de ‘esforço humano’) que é necessário para se validar a inspeção (leitura de placas) com fim de arrecadação tarifária. Não é difícil ver que o tamanho do BackOffice é função do estado da arte da tecnologia de inspeção automática, das condições de tráfego, do estado de conservação da placa, iluminação dentre outras. Portanto, o tamanho do BackOffice é uma função das notas de classificação obtidas.

Além disso, o tamanho desse esforço depende da capacidade que um humano tem em inspecionar imagens de placas – que pode ser medido na forma de tempo médio de inspeção de uma placa. Todos os ingredientes para uma estimativa confiável do trabalho necessário em BackOffice devem ser disponibilizados a partir das notas e tempos de inspeção obtidos.

O esforço do BackOffice é função também do grau de risco existente na decisão de não se inspecionar uma determinada imagem. Uma vez gerado um registro, aqueles que possuem uma leitura OCR válida associada podem ou não ser lidos por um humano. A taxa com que esses registros válidos devem ser lidos não pode ser extraída apenas do valor das taxas e nem de características intrínsecas do sistema. Outros fatores devem ser levados em conta tal como a expectativa de perda de arrecadação ao se decidir não inspecionar registros válidos ou, de outra forma, redução dos custos por se decidir não inspecionar registros potencialmente corretos. Portanto, é necessário considerar as variáveis financeiras nesse dimensionamento.



Figura 73 Processo de validação de um registro em BackOffice.

O objetivo do BackOffice é gerar uma cobrança conforme pode ser visto na Figura 73. Sobre esse processo, já discutimos anteriormente que:

- ▶ Se ocorreu um registro, é obrigatório que se tenha uma imagem;
- ▶ Se se tem uma imagem, essa imagem pode conter uma placa visível;
- ▶ Se há uma placa visível, é possível que essa placa seja legível;

- ▶ Se a placa é legível, é possível que o sistema OCR possa ler a placa e, assim, é possível que haja um código válido;
- ▶ Se há um código válido, é possível que ele seja correto (isto é, refere-se realmente a um veículo que exista na frota);
- ▶ Se o código foi validado, é possível que ele seja realmente a do veículo que fez a passagem.

Portanto, a chance de se ter uma cobrança é o produto das probabilidades intermediárias em todo o processo. No processo de validação da imagem, é possível dimensionar o BackOffice de forma que todas as imagens *que não contenham código válido* sejam inspecionadas ou apenas *uma fração delas* (chamamos de ‘x’ essa fração) o seja. O dimensionamento do BackOffice depende explicitamente de x, pois se  $x = 0$ , não há BackOffice e o erro de validação é proporcional ao erro global com que o sistema de OCR não consegue ler as placas, enquanto que se  $x = 1$ , todas as imagens, necessariamente devem ir para BackOffice a um custo que pode não compensar. A arrecadação total será proporcional a  $\%fCE + x\%fCH$ . O valor de x deve ser calculado a partir de expectativas financeiras que levem em consideração todos os custos envolvidos, valores de pedágio por ponto (eventual dependência do valor com o período de arrecadação), custos fixos (isto é, não apenas com o sistema de arrecadação mas também com o próprio sistema de cobrança, BackOffice, custo associados à execução da cobrança etc). Dadas todas as variáveis envolvidas no problema financeiro, não é difícil assim ver que o custo total de BackOffice não pode ser calculado tão somente a partir dos índices obtidos que pode, no máximo, determinar o esforço total em BackOffice.

Entretanto, podemos estimar o número de pessoas com base em valores diferentes de tempos de inspeção e assumindo que  $x = 1$  *para todos os registros*. Nesse caso, podemos, por exemplo, estimar o número de pessoas trabalhando em BackOffice em regime de turnos de 6 h, 8 h e 10 h para viabilizar a inspeção de todos os registros de todos os sítios conforme uma variação na duração do turno. Para isso, podemos assumir também que não existem perdas de registro (ou seja  $\%IrV = 0$  para todos os sítios, o que está próximo apenas para os sítios de GRU e SJM). As tabelas da Figura 74 e Figura 75 apresentam os valores de pessoas para cada sítio com base em dois tipos de tempo:

- Tempo médio de inspeção da ordem de 4s (conforme calculado na Seção “Tempos de inspeção em BackOffice”);
- Tempo médio de inspeção equivalente a q2 da ordem de 10s. Esse é o tempo que um indivíduo leva para digitar uma placa sem ter que fazer conferência de seu código lido automaticamente. (também estimado na Seção “Tempos de inspeção em BackOffice”).

É bastante provável que a eficiência de inspeção em BackOffice<sup>34</sup> dependa fortemente da duração do turno. Assim, ao se usar turnos mais longos, o número de pessoas envolvidas é menor, mas a eficiência é também menor. Para o tempo de inspeção de 10 s (o sujeito não é obrigado a checar se o código retornado é válido, mas apenas digitar a placa), em um turno de 6h, o número máximo de registros por dia que um indivíduo consegue validar é da ordem de 2160 registros (volume total dividido por 158). Se o turno for de 10h, esse número é 3600 registros a uma eficiência provavelmente reduzida. Observa-se que esse é o número total de pessoas para 1 turno, ao se considerar vários turnos no dia, o número de pessoas trabalhando simultaneamente é inversamente proporcional ao número de turnos.

			TURNOS (pessoas)		
	volumes/dia	Tempo total para média de 4s (horas)	turno 6h	turno 8h	turno 10h
<b>GRU</b>	150000	<b>166,7</b>	27,78	20,83	16,67
<b>SJS</b>	42000	<b>46,7</b>	7,78	5,83	4,67
<b>BMA</b>	21000	<b>23,3</b>	3,89	2,92	2,33
<b>SJM</b>	77000	<b>85,6</b>	14,26	10,69	8,56
<b>SJC</b>	52000	<b>57,8</b>	9,63	7,22	5,78
<b>TOTAL</b>	<b>342000</b>	<b>380,0</b>	<b>63</b>	<b>48</b>	<b>38</b>
<b>Registros/pessoa</b>			<b>5400</b>	<b>7200</b>	<b>9000</b>

Figura 74 Número estimado de pessoas conforme a duração do turno para um tempo médio de 4s por registro. Assume-se que todos os veículos são registrados.

			TURNOS (pessoas)		
	volumes/dia	Tempo total para q2 (10s)	turno 6h	turno 8h	turno 10h
<b>GRU</b>	150000	<b>416,67</b>	69,44	52,08	41,67
<b>SJS</b>	42000	<b>116,67</b>	19,44	14,58	11,67
<b>BMA</b>	21000	<b>58,33</b>	9,72	7,29	5,83
<b>SJM</b>	77000	<b>213,89</b>	35,65	26,74	21,39
<b>SJC</b>	52000	<b>144,44</b>	24,07	18,06	14,44
<b>TOTAL</b>	<b>342000</b>	<b>950,0</b>	<b>158</b>	<b>119</b>	<b>95</b>
<b>Registros/pessoa</b>			<b>2160</b>	<b>2880</b>	<b>3600</b>

Figura 75 Número estimado de pessoas conforme a duração do turno para um tempo médio de 10s por registro (tempo médio para classe q2). Assume-se que todos os veículos são registrados.

<sup>34</sup> Isto é, a taxa de acerto com que humanos conseguem validar placas.

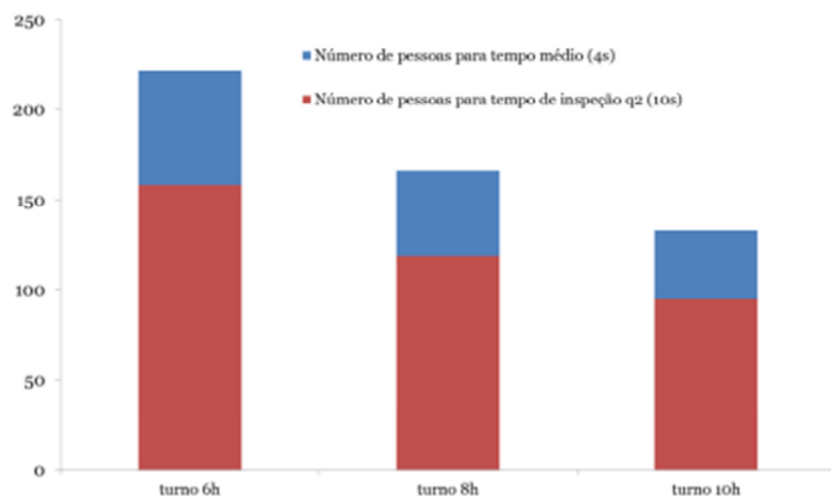


Figura 76 Comparação do número de pessoas necessárias para inspeção de todos os registros dos 5 sítios para diversas durações de turno. Esse gráfico corresponde aos totais mostrados nas duas figuras anteriores.

## Índices de rastreabilidade de veículos potencial do sistema

A rastreabilidade implica em ser possível determinar o tempo de trajeto de um determinado veículo ao menos entre 2 sítios de registro. Um teste de rastreabilidade pode ser feito pelos pontos de aquisição do projeto de duas formas diferentes:

- Buscas exaustivas em todos os dados gerados para detecção de veículos com placas validadas correspondentes a cada passagem<sup>35</sup>. Essa busca teria que ser feita para todos os dados gerados *que deveriam ser totalmente inspecionados*, o que foge do escopo do projeto, que é a estimativa estatística (por amostragem) do desempenho de leitura automática de placas. Além disso, não se tem certeza que, dado um veículo, ele, de fato, cumpriu todo o itinerário que passa por pelo menos 4 pontos (do Rio de Janeiro a São Paulo);
- De posse de um grupo de veículos conhecidos, executar a passagem do grupo por todos os pontos, com registro dos horários de passagem seguindo o horário sincronizado interno do sistema. Com isso seria possível buscar de forma facilitada o registro da passagem de cada veículo sem ter que explorar os registros. Depois disso, seria feita a inspeção das placas capturadas e calculado um índice específico como ‘chance de rastreabilidade’ do sistema. Para que a estatística seja representativa, seria necessário escolher um número apropriado de veículos (no mínimo 100) para o cálculo desses índices.

<sup>35</sup> Um veículo teria que ser buscado em todos os registros validados dos outros pontos sucessivamente. O projeto não teve tempo suficiente para realizar essa análise, que exigiria inspeção humana de todos os registros e compilação em banco de dados para executar a busca de forma automática.

Há, porém uma estimativa teórica que podemos fazer facilmente, conhecidas as notas individuais de cada ponto, isto é, a chance de se validar uma placa corretamente (independente de ser sistema automático ou humano) em um dado ponto. É possível determinar ‘índices de rastreabilidade’ ou capacidade de se rastrear um determinado veículo para um número arbitrário de sítios conforme as frações de acerto obtidas. A possibilidade de rastreabilidade entre M sítios só está garantida para veículos que efetivamente atravessaram cada um dos M sítios. A chance de se rastrear um veículo por M pórticos,  $P_{RM}$ , será dada por:

$$P_{RM} = \prod_k (\%fCE_k + \%fCH_k),$$

onde o produto é feito sobre todos os M pórticos. Nesse cálculo, a chance de se ler corretamente uma placa é dada pela soma da fração de leitura automática com a fração de leitura do humano. O erro associado a essa chance de rastreabilidade será dado por

$$\sigma_{RM} = \sum_j \prod_{k \neq j} (\%fCE_k + \%fCH_k) \sigma_k$$

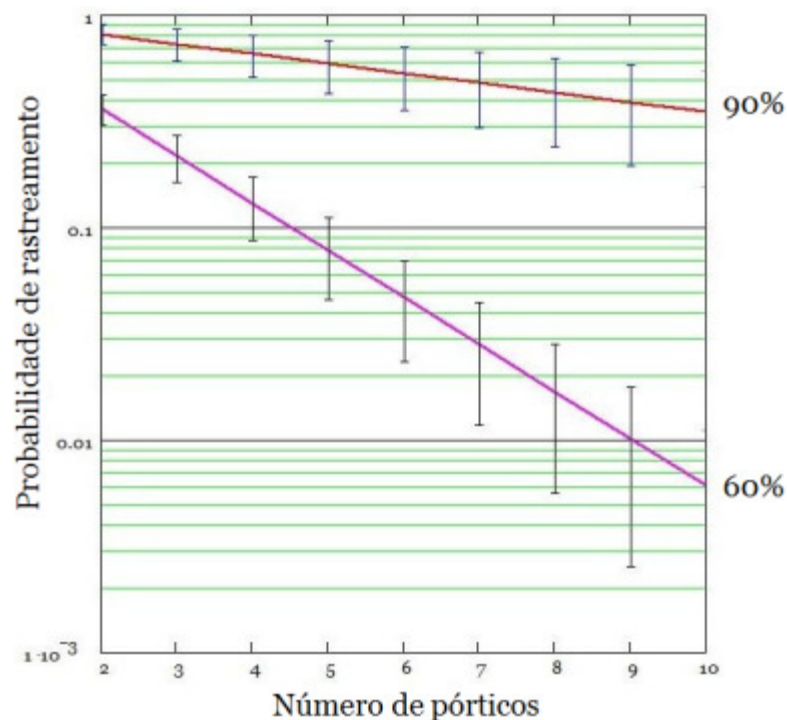


Figura 77 Probabilidades de captura como função do número de pórticos (pelo menos 2) e probabilidade global de leitura correta única (soma das notas de acerto automática e humana), 90% (vermelho) e 60% (lilás). O erro associado corresponde a um desvio padrão de 5%

onde a soma é feita sobre todos os pórticos e o produto exclui sucessivamente cada pórtico onde o erro é considerado. Por exemplo, se  $\%fCE + \%fCH = 90\% \pm 5\%$  (que é

um valor aproximadamente igual as melhores notas obtidas durante o projeto admitindo %IrV=0) então temos:

$$P_{R2}=81,0\%\pm9,0\%$$

$$P_{R3}=72,9\%\pm12,1\%$$

$$P_{R4}=65,6\%\pm14,5\%$$

Portanto, a chance de um veículo ser rastreado em 2 pórticos é igual a 81% para essa probabilidade de acerto único de 90%. À medida que as probabilidades de leitura correta de cada pórtico vão diminuindo, a chance de rastreamento também cai e o erro associado aumenta. A chance média de se capturar um veículo em 4 pórticos poderá estar entre 51,1% e 80,1%. De modo geral, a chance de se conseguir rastrear um veículo sucessivamente por 2 até 10 pórticos pode ser vista na Figura 77 juntamente com as expectativas de erro para 90% (curva vermelha) e 60% (curva lilás) e um nível de 5% de erro. Para o nível de acerto único de 60%, a partir do uso de 4 pontos de aferição, a probabilidade de rastreamento é inferior a 10%.



## Conclusões

Análises a guisa de conclusão são dadas abaixo:

1. O método implementado para inspeção em BackOffice demonstrou sucesso na inspeção para geração de estatísticas das diversas notas pré-definidas conforme a Seção “Hierarquia binomial usada e notas finais”. Essas notas podem ser usadas de forma ostensiva na caracterização da excelência de inspeção de qualquer sistema de inspeção de placas. Em particular, foi visto que os índices %fPAL e %fIL (percentual de placas assertivamente válidas e percentual de registros que retornam código válido respectivamente), embora representando um nível de acesso à qualidade final do sistema de leitura OCR, não são indicadores globalmente válidos. Em seu lugar, deve-se considerar o uso de %fCE, %fII e %fCH em conjunto com %IrV como índices de caracterização apropriados do sistema;
2. Em particular, o resultado final das notas teve que ser corrigido para os fluxos aferidos no momento de passagem por um sistema independente da contagem de tráfego. Por meio da criação de um índice adicional, %IrV, foi possível conhecer o valor real de cada nota. É importante considerar que as perdas totais são obtidas da soma de diversos termos. Para cada sítio  $j$  e período  $k$ , a perda total será dada por  $\%fPI(j,k) + \%fPP(j,k) + \%IrV(j,k)$ . Idealmente,  $\%IrV(j,k) \rightarrow 0$  com os sistemas mais bem ajustados, de forma que a perda total é equivalente a  $\%fII(j,k) = \%fPI(j,k) + \%fPP(j,k)$ . Os sistemas instalados em rodovia, ainda que tenham uma chance não nula de perder veículos, devem necessariamente registrar essa perda em sistema. Durante o projeto, foi exigido que os fornecedores produzissem, em tempo real, *relatórios de log de sistema*, que são relatórios contendo instruções ou descrições internas de cada sistema de câmera sobre seu status de funcionamento, paradas, eventuais perdas de comunicação etc.
3. Embora não tenha sido apresentado nas seções anteriores, apenas o sítio GRU produziu relatórios de log suficientemente completos que registram também perdas sistêmicas de captura. Assim, ainda que se observem perdas pela análise de %IrV, o sistema de GRU reportou estar ciente disso. Esses relatórios são importantes e suplementares para a caracterização da solução utilizada e todos os registros produzidos em um determinado ponto de inspeção, já que representam a ‘vida real’ e não o que fornecedores em geral idealizam com relação a seus sistemas. Acreditamos que todos outros fornecedores devem, em tempo, considerar a produção desses relatórios completos, que serão fundamentais para a contextualização dos registros em BackOffice;
4. Espera-se internamente que o número de registros seja igual ao número de imagens. Entretanto, problemas internos podem fazer com que o número de

imagens seja menor (por conta de perdas ou dificuldades de gravação da imagem), o que é esperado. Entretanto, o número de imagens *não pode ser maior*. Em alguns pontos (SJM e SJC) foi observado que imagens foram geradas em excesso ao número de registros. Isso se constitui um problema grave já que a inspeção em BackOffice despreza naturalmente (por inexistência de referência de registro) as imagens em excesso. Porém, em SJM foi constatado que %IrV tornou-se negativo em alguns períodos (veja Figura 36), por conta de geração de registros em excesso ao número real de veículos. Se esses registros estão associados a imagens, isso pode gerar trabalho adicional em BackOffice. É importante considerar que o uso de registros e imagens múltiplas pode ser uma decisão de estratégia de inspeção – aumento de redundância – mas está necessariamente associado a trabalho adicional de inspeção;

5. As notas ou índices observados ao longo de vários períodos em dois dias escolhidos de funcionamento comum de todos os sítios demonstram uma variedade *inter* sítios e uma variação *intra* sítios. A variação *inter* sítios foi maior que a esperada originalmente<sup>36</sup>. Com relação à variação *intra* sítios, as curvas para GRU e SJS são razoavelmente homogêneas em um longo período do dia (considerar %fCE e %fIL na Figura 61, Figura 62, Figura 63 e Figura 64) apresentando quedas nos períodos noturnos. Essas quedas estão correlacionadas com acréscimo do percentual %fCH, indicando requisição de BackOffice na inspeção de imagens não conformes. Isso é totalmente esperado segundo o estado da arte da tecnologia de OCR. Assim, a curva de desempenho de sistemas OCR apresenta máximos nos períodos de maior luminosidade (meio do dia) e mínimos no período noturno. Tal observação é importante para o planejamento do BackOffice já que esse comitê será menos requisitado nos períodos de maior luminosidade (taxa maior de registros válidos);

---

<sup>36</sup> Por conta de fatores ligados à tecnologia ou falha de configuração de equipamentos.

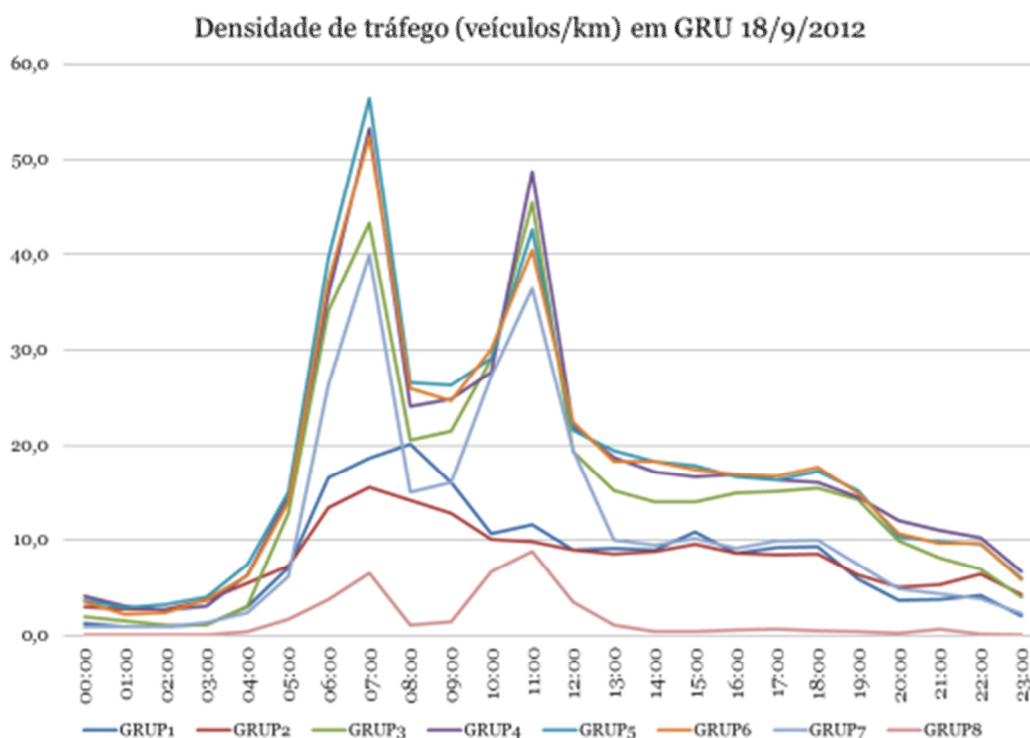


Figura 78 Densidade de tráfego como função do período do dia em GRU em 18/9/2012 para as diversas pistas do local.

6. O comportamento dos sistemas GRU e SJS convergidos apenas demonstram correlação com a luminosidade. Não se constatou, durante o projeto, qualquer correlação das curvas com a *densidade de tráfego*. A Figura 78 é um gráfico dessa densidade como função da hora para diversas pistas no sítio GRU em 18/9/2012, em um dos dias que fazem parte do período oficial. É possível ver dois picos de densidade em torno das 6:00 e 11:00 para todas as pistas (exceto pela P8 que é praticamente um acostamento) e que indicam presença de congestionamento. Se tais períodos são inspecionados na Figura 62, observamos que não houve modificação alguma nas notas de desempenho. Isso indica que, para o nível de congestionamento da ordem de 50 a 60 veículos por km, o sistema de inspeção instalado em GRU é robusto e apenas dependente do nível de iluminação como esperado;
7. Ainda com relação às perdas intrínsecas de registro, notamos que as perdas maiores ocorreram em sistemas que não utilizam processos ativos de detecção. Um processo ativo de detecção faz uso de instrumental de pista que é acionado com a passagem fisicamente próxima do veículo. Sistemas passivos ou ‘virtuais’ tais como os chamados ‘laços virtuais’ dependem de variações consideráveis de iluminação externa para que sejam efetivos. Embora acreditemos que esses sistemas são, de fato, eficientes, eles exigem condições mais rigorosas para seu funcionamento pleno. Dentre elas, a mais importante é garantir um nível mínimo de iluminação externo que seja comensurável com a sensibilidade das câmeras que dele fazem parte. Isso significa que, em tais sistemas, iluminação externa é necessária, o que parece indicar que o uso de

laços virtuais é propício a ambientes urbanos onde iluminação pública já está disponível. Em ambientes rurais, postes com iluminação adicional devem ser utilizados. Embora não seja escopo do projeto, é provável que iluminação infravermelha seja suficiente para a função de contraste. De qualquer forma, laços virtuais, se por um lado apresentam a vantagem de não exigir intervenções na pista, tem a desvantagem de exigir iluminação. É provável, além disso, que mesmo com iluminação adicional, certas cores de veículos influenciem na taxa global de aquisição, o que deve ser confirmado por testes adicionais;

8. Por outro lado, as curvas de %fCE observadas nos sítios de BMA, SJM e SJC (Figura 65 a Figura 70) são diferentes das curvas equivalentes para GRU e SJS, o que caracteriza a variação *inter* sítios. Isso indica que os sistemas de BMA, SJM e SJC não estão ajustados o suficiente para a tarefa de registro e inspeção de imagens. Além disso, as notas finais para esses sítios – corrigidas segundo %IrV – resultaram muito baixas por conta da perda de veículos. Em particular, o sítio de SJM, embora apresente o melhor desempenho de todos no registro de veículos, está com as notas de desempenho em OCR inerentemente baixas, apresentando períodos de geração espúria de registros conforme comentado anteriormente. As curvas de desempenho %fCE em SJC (Figura 69 e Figura 70) estão abaixo e coincidentes com a curva %fPAL. Isso acontece porque em SJC apenas os registros que resultaram em códigos válidos estão sendo reportados (o mesmo ocorre em BMA). Pela definição de %fPAL e %fCE, se o número de imagens com registros válidos ( $m$ ) for igual ao número total de registros ( $n$ ), as duas notas coincidem;
9. Com relação à estatística de motos (índice %fT), foi necessário segregar as imagens que correspondem a esse tipo de veículo. Essa foi, na verdade, o único tipo de classificação veicular que foi feito. Conforme mostram as tabelas da Figura 71, os sítios de GRU e SJS foram os únicos que registraram presença de motos (há uma pequena contribuição residual em SJM em 18/9 que é provavelmente espúria) e a estatística de presença indica taxas médias da ordem de  $4,18\% \pm 0,33\%$  e  $3,68\% \pm 0,31\%$  em 17/9 e 18/9, respectivamente, em GRU e  $3,38\% \pm 0,45\%$  e  $2,19\% \pm 0,33\%$  em 17/9 e 18/9, respectivamente, em SJS. Portanto, a quantidade de motos parece estar limitada a menos de 5% do fluxo total. Ainda assim, o sistema instalado em SJS (que capturava placas traseiras de veículos) conseguiu detectar e retornar códigos validados para aproximadamente  $\frac{1}{4}$  do volume de motos. A inspeção humana torna-se especialmente necessária na maioria dos casos (cerca de 54,7%) e aproximadamente 29% do volume é perdido;
10. Conforme descrito na Seção “Estatísticas específicas”, não foi detectada variação apreciável dos índices de inspeção com a incidência de chuva com base em dados de GRU. O período escolhido (5:00) corresponde a um período de baixa iluminação e tráfego pequeno (ver Figura 78) que antecede a um período de congestionamento;

11. Com relação ao planejamento do BackOffice em uma operação real do sistema, a análise dos tempos de inspeção (propiciado pela medida desses tempos por meio de uma interface especial criada durante o projeto) determinou um tempo médio de 4s (ver Seção “Dimensão estimada de BackOffice”). Não obstante esse valor, é importante considerar que o tempo real de inspeção por registro dependerá bastante da maneira como serão organizados os dados e sua distribuição entre diversos inspetores. Há várias estratégias de inspeção, sendo que a frequência com que se inspeciona uma imagem depende de seu status no sistema (se ela apresenta um código válido é apenas uma dessas caracterizações). O fato de o código válido apresentar certa sequência de letras ou o registro ter sido obtido em determinadas horas propícias ou não do dia são outras caracterizações possíveis. Portanto, o tempo efetivo de inspeção dependerá bastante desses arranjos especiais, e aqui apenas determinamos medidas aproximadas de esforço de BackOffice conforme o tempo médio ou o tempo médio de digitação (aproximadamente 10 s), na hipótese de todos os registros serem inspecionados. Em particular, nesse último caso, um turno de 6 horas de inspeção determina um máximo de aproximadamente 2200 registros/dia por inspetor, enquanto que valores progressivamente maiores são obtidos com turnos mais longos. Chamamos a atenção que essa ‘duração de turno’ é o *tempo líquido ou efetivo de inspeção*. Assim, um turno de 6h pode corresponder a um turno oficial de maior período, conforme a política adotada de descanso ou horários livres (almoço, intervalos de lanche, descanso etc). Considerado esse intervalo de 6h, um número máximo de 160 pessoas seria necessário para inspecionar todo o fluxo diário pelos 5 sítios que fizeram parte do projeto, onde consideramos os sistemas ajustados e sem perda em relação ao tráfego real. Essa análise deve ser refeita ao se considerar a necessidade de inspeção de imagens da pista sentido ‘norte’ num total de 8 sítios e um fluxo aproximadamente dobrado (a menos de SJC, o fluxo em um sentido deve considerar apenas os volumes aferidos em GRU, SJS, BMA e SJM);
12. No que diz respeito à rastreabilidade, a seção “Índices de rastreabilidade de veículos potencial do sistema” apresentou algumas estimativas teóricas sobre a chance de se rastrear um veículo ao longo de uma sequência de sítios. Nesse cálculo, novamente, podemos assumir que não há perdas de tráfego ( $\%IrV=0$ ), e a nota máxima é obtida fazendo-se  $\%fCE + \%fCH$ . É possível ver que as chances de rastreamento são superiores a 50% até um máximo de 7 pórticos para as máximas taxas de validação. O projeto não executou teste de rastreamento real nem busca no banco de registro por razões que são discutidas na seção citada.

### Sugestões para estudos futuros

Dentre as sugestões de estudos complementares futuros com o sistema implementado, citamos:

- Proposta de continuação da aquisição com resolução mais acurada do nível de tráfego. Para todos os pontos em que se observam falhas, aprimorar o registro e gravação de imagens de forma que elas fiquem dentro do esperado para perda de tráfego em relação ao sistema de aferição e eliminação de instâncias de multiplicidade de registros;
- Adquirir sequência de intervalos contínuos em período de chuvas intermitentes. Como já foi dito, o projeto se desenvolveu ao longo de um período de forte estiagem. Os meses de Dezembro a Fevereiro são reconhecidamente os mais chuvosos. A escolha de um período contínuo de aquisição, por exemplo, Janeiro, pode possibilitar o estudo mais sistemático do nível de influência da chuva ao longo de vários períodos do dia e não apenas à noite. A amostragem a ser usada é a mesma que foi feita para os dias de 17/9/2012 e 18/9/2012. Idealmente essa análise deveria ser feita com o auxílio de ferramental de medida de chuva (pluviômetro) instalado em cada sítio para que os períodos de chuva fossem precisamente determinados;
- Implementação de processo de classificação veicular e vinculação da nota ao tipo de veículo. Dentro os vários segmentos utilizados para a determinação das diversas frações, a posição da aquisição e o período foram determinados. É possível também segregar essas notas por tipo de veículo, usando a imagem frontal ou traseira como parâmetro de classificação. Para isso, além da inspeção humana da placa, seria requisitado que o operador de BackOffice também selecionasse um conjunto de botões com o tipo de veículo: veículo de passeio, caminhão, ônibus e moto. Isso poderia ser feito tanto com imagens da dianteira como traseira de veículos. A determinação das notas por tipo de veículo em BackOffice naturalmente aumenta o período de inspeção e está sujeita a erros, mas pode resultar em estimativas de variação das notas segundo esse critério limitado de classificação;
- De certa forma, o sítio de GRU já possui um sistema de segregação de tipo de veículo por comprimento e os dados fornecidos já incluem esse dado. Um estudo que pode ser feito com os dados de GRU é segregar as notas de acordo com classes de comprimento. Exemplo: veículos de 0 a 4 m, veículos de 4 a 10 metros, veículos de 10 a 20 metros, veículos acima de 20 m etc. Para isso, os volumes de veículo por comprimento seriam definidos de acordo com o tipo de pista e um esquema de amostragem incluindo mais veículos seria utilizado. A inspeção em BackOffice seria refeita visando incluir mais veículos segregados por comprimento segundo esse esquema. Um estudo desse tipo teria a vantagem de naturalmente determinar as frações por tipo de placa associado ao comprimento, capturando variações de contexto da placa e seu posicionamento;
- Outra análise possível ainda com os dados já disponíveis é a segregação por velocidade. Essa análise somente poderia ser feita no sítio GRU que dispõe de recurso para calcular a velocidade instantânea. Assim, seria possível investigar o efeito da velocidade sobre a eficiência de inspeção (essencialmente sobre %fCE e %fCH) para um intervalo de velocidades (p. ex., de 60km/h a



140km/h). É esperada uma degradação na qualidade da imagem com velocidades maiores (redução no tempo de exposição com redução do contraste), e a análise poderia ser feita apenas para veículos que apresentarem excesso;

- Determinação acurada das classes de letras onde ocorrem maiores erros. Da inspeção em BackOffice é aparente que placas com determinadas letras têm maior frequência de erros. O levantamento dos percentuais relativos no alfabeto para cada letra é importante para as estratégias de inspeção em BackOffice, onde placas contendo esses agrupamentos de letras seriam obrigatoriamente inspecionadas, mesmo que resultassem em códigos válidos para os melhores períodos e condições de aquisição. Essa estatística é independente das condições de aquisição (não depende de clima, condições de iluminação etc) e, portanto, pode ser feita com o grupo de dados já adquiridos.

## **Apêndice 1 – Formatação de saída de dados para análise em BackOffice do projeto “Sistema Automático de Arrecadação para Teste”**

O objetivo deste documento é prover informações sobre a organização de diretórios, registro e identificação de fontes de dados para o projeto “Sistema Automático de Arrecadação para Teste”, a ser realizado na rodovia Nova Dutra, sob a organização da Engellog/CCR (Sistema Automático de Arrecadação para Teste, 2012).

### **Transmissão dos arquivos**

A transferência dos diretórios e arquivos ao Servidor do centro de Dados e deverá ocorrer imediatamente após a finalização (fechamento) dos arquivos a serem transferidos; ou seja, os arquivos de imagens deverão imediatamente ser enviados, e os arquivos de dados enviados de 30 em 30 minutos (ao final de cada intervalo), e o arquivo de LOG devem ser enviados diariamente (ao final de cada dia).



## Organização de Diretórios

### 1.1. Diretório Raiz

É o diretório que contém todos os demais diretórios de cada pórtilco, e que será replicado em sua íntegra no servidor do Centro de Dados.

Formato:

<ID\_LOCAL>

As siglas identificadoras dos locais (ID\_LOCAL) encontram-se especificadas na Tabela 3a.

Exemplo:

..\GRU\

Campo	Código	Descrição	Tamanho	Tipo original
1	ID_LOCAL	Identificação do pórtilco	3	Alfabético

### 1.2. Diretório Diários

São os diretórios separados por dia (das 00h00m00s até 23h59m59,999s) e terão a seguinte nomenclatura:

Formato:

<ID\_LOCAL>\_<AAAA>\_<MM>\_<DD>

As siglas identificadoras dos locais (ID\_LOCAL) encontram-se especificadas na Tabela 3a.

Exemplo:

..\GRU\_2012\_05\_20\

Tabela 1a. Campos para geração do nome do diretório principal.

Campo	Código	Descrição	Tamanho	Tipo original
1	ID_LOCAL	Identificação do pórtilco	3	Alfabético

2	AAAA	Ano	4	Numérico inteiro
3	MM	Mês	2	Numérico inteiro
4	DD	Dia	2	Numérico inteiro

### 1.3. Subdiretório Pista

Dentro de cada diretório principal, haverá uma separação por faixa, dada pela letra “P” seguida pelo número da faixa de rolagem. O número da faixa deverá ser contado da esquerda (faixa de maior velocidade) para a direita. Havendo mais de uma pista no mesmo sentido, sempre começar a contar as faixas pela pista da esquerda, ou seja, a pista de maior velocidade é a 1, conforme indicado na Figura A, com três faixas.

Formato:

P<NÚMERO DA PISTA>

Exemplo: Para a Pista 3, tem-se:

..\P3\

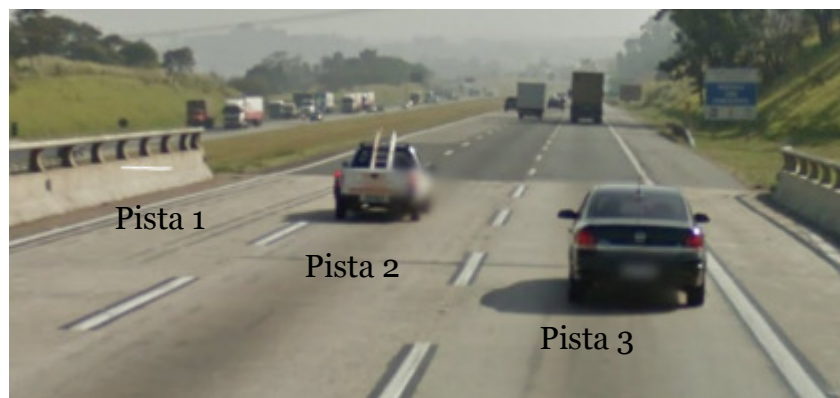


Figura A. Exemplo de identificação das pistas para um caso com três faixas (pistas) em um dado sentido de tráfego.

### 1.4. Subdiretórios de dados e imagens

As imagens correspondentes aos arquivos de dados serão armazenadas em diretórios com a mesma nomenclatura do arquivo TXT do arquivo de dados.

O nome do diretório de imagens terá a seguinte formação:

Formato:

<ID\_LOCAL>\_<AAAA>\_<MM>\_<DD>\_<INTERVALO>\_<NÚMERO DA PISTA>, onde o INTERVALO deve ser calculado em função do horário de captura da imagem (passagem do veículo), de acordo com o mapeamento apresentado na Tabela 4a.

Exemplo:

..\GRU\_2012\_05\_22\_01\_P1\ -> corresponde a dados coletados no pórtico de Guarulhos, intervalo 01 (das 00:00:00,000 às 00:29:59,999) e Pista 1.

### 1.5.Exemplo de construção de diretórios

- Contexto de operação: suponha um caso ilustrativo em que foram capturadas imagens ao longo de 3 meses, todos os dias. Suponha os meses de maio (05), junho (06) e agosto (08). A estrutura de diretórios para o pórtico de Guarulhos é exibida na Figura B. As estruturas de diretórios para os pórticos de São José dos Campos, São José do Meriti e Barra Mansa são exibidas nas Figuras C, D e E, respectivamente.

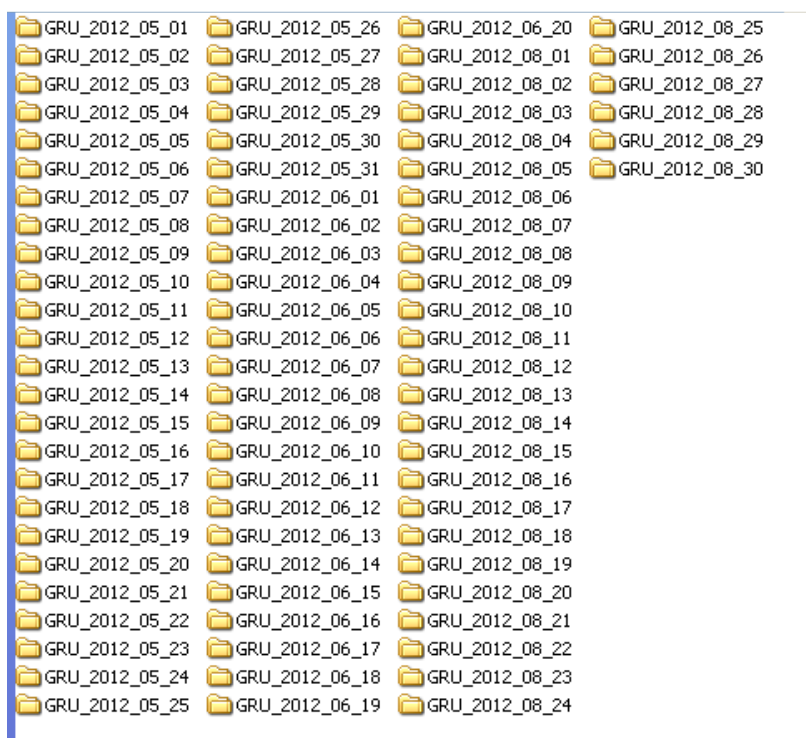


Figura B. Exemplo da estrutura principal de diretórios, localizados um nível abaixo do diretório principal \GRU, para sistema operante todos os dias ao longo dos meses de maio, junho e agosto de 2012 no Pórtico de Guarulhos.

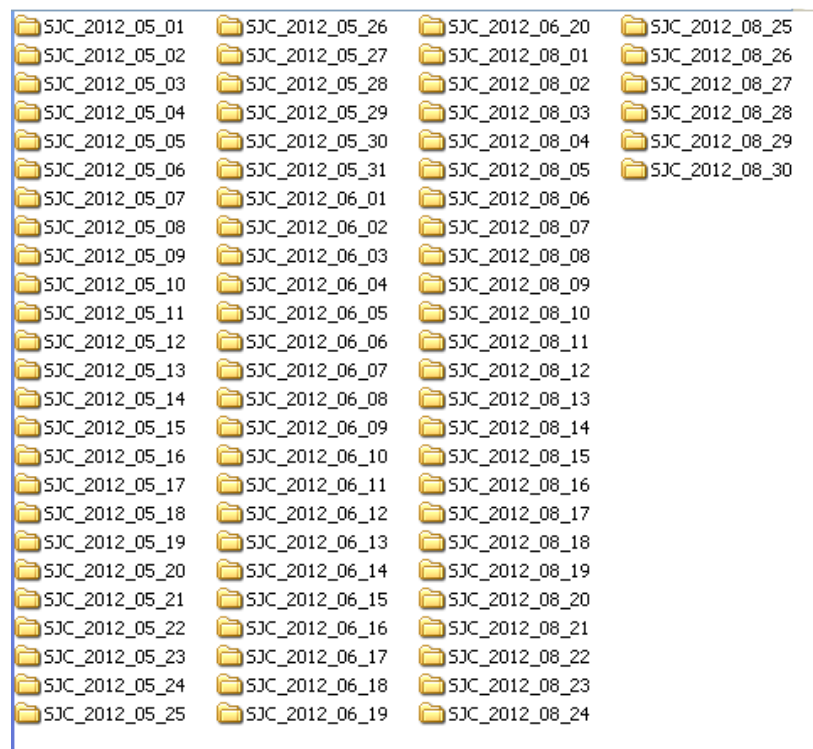


Figura C. Exemplo da estrutura de diretórios, localizados um nível abaixo do diretório principal \SJC, para sistema operante todos os dias ao longo dos meses de maio, junho e agosto de 2012 no Pórtico de São José dos Campos.

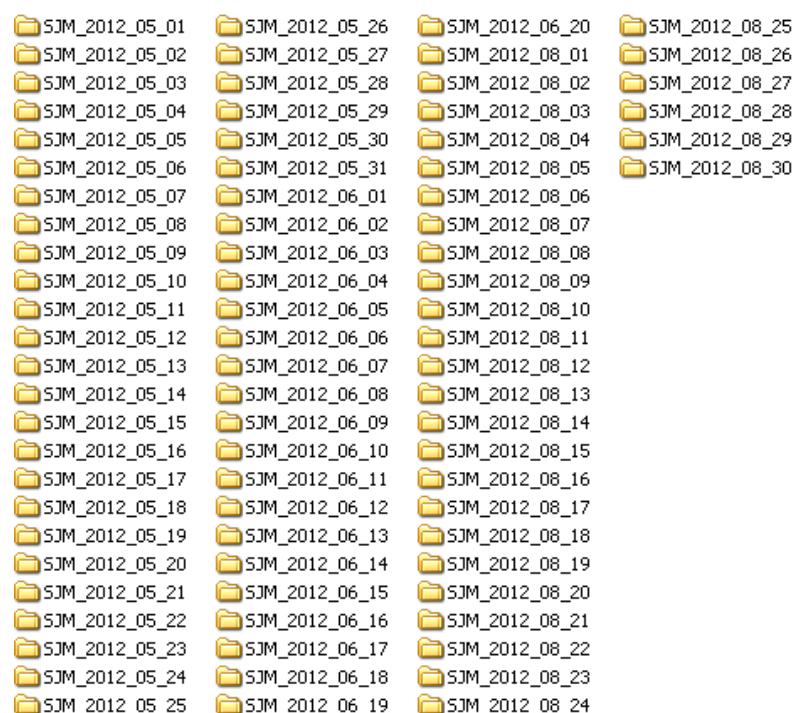


Figura D. Exemplo da estrutura principal de diretórios, localizados um nível abaixo do diretório principal \SJM, para sistema operante todos os dias ao longo dos meses de maio, junho e agosto de 2012 no Pórtico de São José dos Meriti.

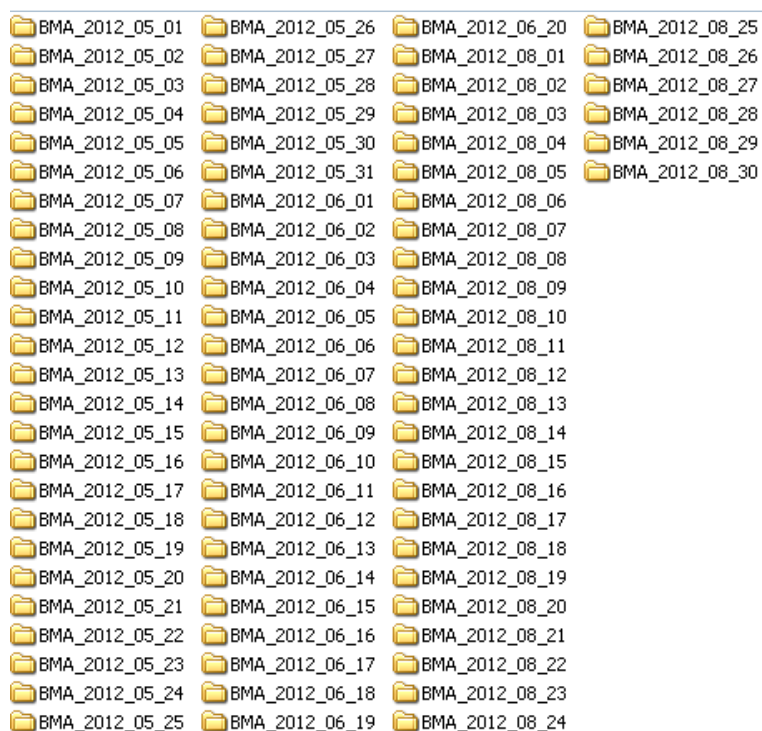


Figura E. Exemplo da estrutura principal de diretórios, localizados um nível abaixo do diretório principal \BMA, para sistema operante todos os dias ao longo dos meses de maio, junho e agosto de 2012 no Pórtico de Barra Mansa.

- Dentro de cada diretório principal, haverá um para cada pista. Supondo 4 pistas, a Figura F exibe a estrutura de um dos diretórios principais qualquer.

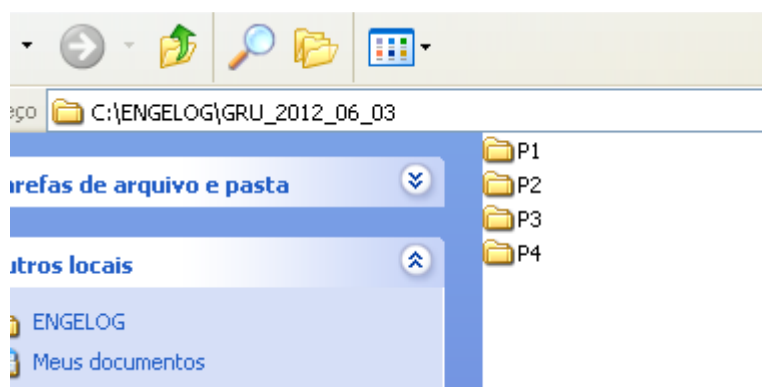


Figura F. Um nível de diretório acima dos principais, há um subdiretório para cada pista. No caso deste exemplo, há 4 pistas.

- Consideremos agora o dia 30/05/2012. Todas as imagens e arquivos de dados coletados neste dia estarão no diretório GRU\_2012\_05\_30. Entrando neste diretório, teremos a configuração exibida na Figura G.

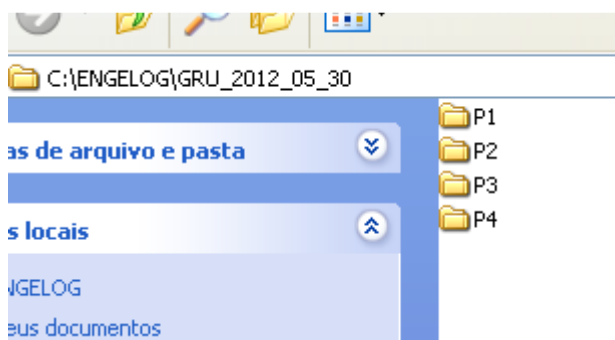


Figura G. Estrutura de subdiretórios para o dia 30/05/2012, pórtico de Guarulhos.

- Supondo agora que queiramos visualizar os arquivos coletados para a Pista 1. Neste caso, entramos no subdiretório “P1”. A Figura H mostra a lista de arquivos e diretórios. No caso, o arquivo “GRU\_2012\_05\_30\_05\_P1.txt” contém os dados coletados dos veículos que passaram pelo Pórtico de Guarulhos no dia 30/05/2012 na Pista 1, e no intervalo de horário 05 (das 02:00:00,000 às 02:29:59,000). O diretório correspondente, denominado “GRU\_2012\_05\_30\_05\_P1”, contém todas as imagens dos veículos relacionados no arquivo texto. Tais imagens não possuem regra de nomenclatura, apenas necessitam ter seu nome informado no campo especificado do arquivo de dados.

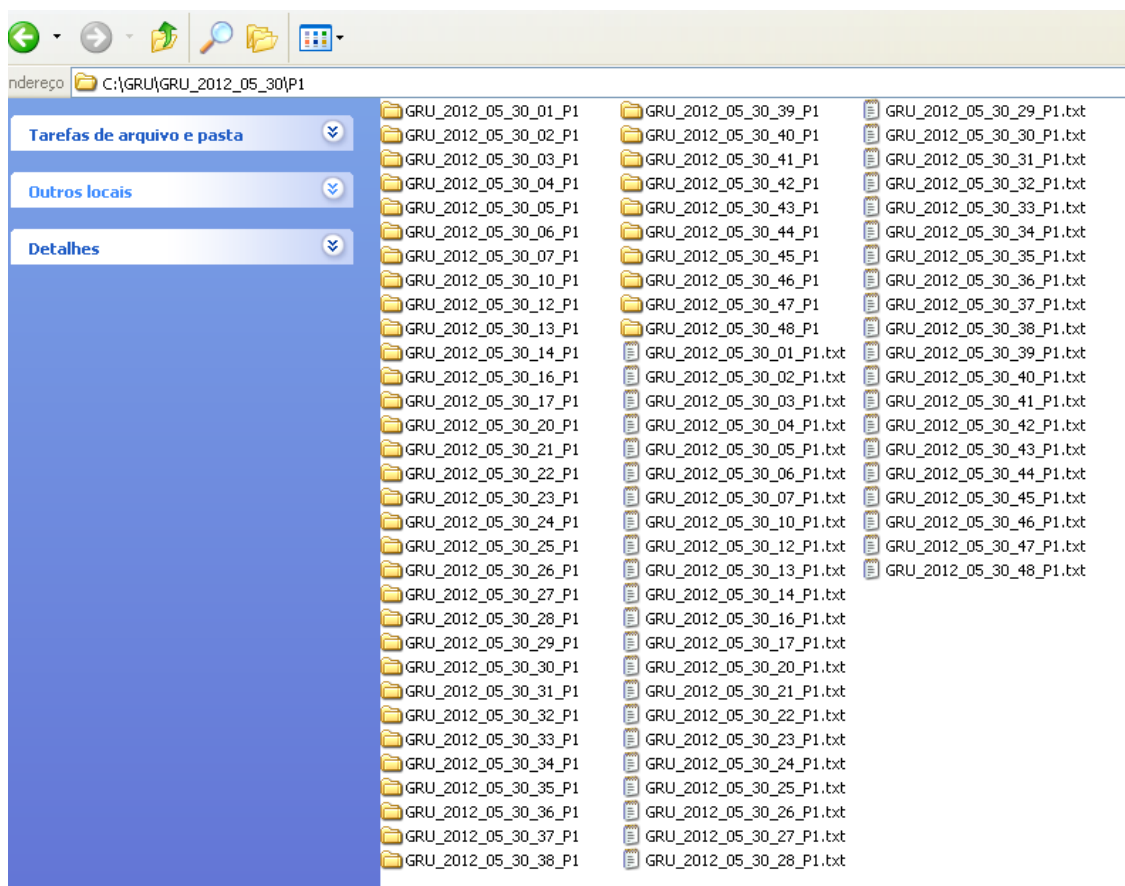


Figura H. Estrutura de arquivos e subdiretórios dentro da pista 1, dia 30/05/2012, pátio de Guarulhos. Aqui contém arquivos referentes a veículos coletados nas 48 faixas de horário. Os diretórios com o mesmo nome dos arquivos remetem às imagens correspondentes.

## Arquivo de Dados

Os dados obtidos do sistema LAP para uma determinada faixa de rolagem serão armazenados em arquivos de dados TXT contendo um intervalo de 30 minutos. Portanto, será gerado um novo arquivo de dados TXT a cada intervalo de tempo, contendo múltiplas passagens numa mesma faixa.

O nome do arquivo de dados terá a seguinte formação:

Formato:

<ID\_LOCAL>\_<DATA>\_<INTERVALO>\_<PISTA>.txt

Em que a data é definida no formato AAAA/MM/DD

Exemplo:

GRU\_2012\_05\_30\_01\_P1.txt

Refere-se a dados coletados no pátio de Guarulhos, no dia 30/05/2012, intervalo de horário 1 e Pista 1.

## Registro de Dados

Cada registro contido no arquivo de dados terá o seguinte formato:

Formato:

<ID\_LOCAL>;<AAAA>;<MM>;<DD>;<HH>;<MM>;<SS>;<mmmm>;<XXX9999>;<GGG>;<CCC>;<F>;<ILU>;<COMP>;<VEL>;<NOME\_IMG>;<CAMPO RESERVADO 1>;<NOME CAMPO RESERVADO 1>;<CAMPO RESERVADO 2>;<NOME CAMPO RESERVADO 2>;<CAMPO RESERVADO 3>;<NOME CAMPO RESERVADO 3>;<CAMPO RESERVADO 4>;<NOME CAMPO RESERVADO 4>;<CAMPO RESERVADO 5>;<NOME CAMPO RESERVADO 5>;<OBSERVAÇÕES>

Exemplo:

A seguir é mostrado um exemplo de uma linha de arquivo, correspondente à captura de uma imagem de veículo. Neste caso, foi utilizado apenas um dos campos reservados (Campo reservado 1), correspondente ao Número de eixos do veículo. Os demais campos reservados não foram utilizados. Como observação (último campo) foi inserido “Dados coletados normalmente”.

GRU;2012;05;20;13;41;32;5221;XYZ1234;95;85;3;DES;650;100;nomeimagem.png;2;  
;



Numero de eixos,,,,,,,,; Dados coletados normalmente

Tabela 1a. Campos que devem conter o arquivo-texto gerado para cada passagem de veículo.

#	Campo	Descrição	Número de caracteres	Tipo original
<b>1</b>	ID_LOCAL	Identificação do fornecedor	3	Alfabético Não pode conter o caractere ‘;’
<b>2</b>	AAAA	Ano	4	Numérico inteiro
<b>3</b>	MM	Mês	2	Numérico inteiro
<b>4</b>	DD	Dia	2	Numérico inteiro
<b>5</b>	HH	Hora	2	Numérico inteiro
<b>6</b>	MM	Minutos	2	Numérico inteiro
<b>7</b>	SS	Segundos	2	Numérico inteiro
<b>8</b>	mmm	Milissegundos	3	Numérico inteiro
<b>9</b>	XXX9999	Placa	7	Alfanumérico, podendo conter somente letras e dígitos
<b>10</b>	GGG	Taxa de confiabilidade de acerto global da leitura da placa, em porcentagem	3	Numérico inteiro
<b>11</b>	CCC	Menor taxa de confiabilidade dentre todos os caracteres da placa, em porcentagem.	3	Numérico inteiro
<b>12</b>	F	Faixa de rodagem de tráfego (número da faixa, da esquerda para a direita)	1	Numérico inteiro
<b>13</b>	ILU	Status da iluminação artificial (LIG/DES)	3	Alfabético
<b>14</b>	COMP	Comprimento aproximado do veículo (em centímetros) no formato “9999”, valor inteiro	4	Numérico inteiro
<b>15</b>	VEL	Velocidade do veículo em km/h, valor inteiro. Quando houver mais de um dado de velocidade, utilizar a velocidade média ou aferida.	3	Numérico inteiro

<b>16</b>	NOME_IMG	Nome da imagem correspondente, conforme a regra de formação de nomes descrita neste documento.	≤ 255	Alfanumérico  Não pode conter o caractere ‘;’
<b>17</b>	SEQ_NUM	Número sequencial por faixa de rodagem, referente ao número da imagem disparada pela câmera.	7	Numérico longo inteiro ( <i>long int</i> )
<b>18</b>	SN_EQUIPTO	Identificação única (serial number) do equipamento no qual foi gerado o log. Considera-se um equipamento (câmera) por pista.	≤ 255	Alfabético  Não pode conter o caractere ‘;’
<b>19</b>	CAMPO RESERVADO 1	Valor do campo dedicado ao armazenamento de dados não solicitados, mas que o fornecedor eventualmente disponha (ex. número de eixos)	≤ 255	Alfanumérico  Não pode conter o caractere ‘;’
<b>20</b>	NOME CAMPO RESERVADO 2	Nome do campo	≤ 255	Alfanumérico  Não pode conter o caractere ‘;’
<b>21</b>	CAMPO RESERVADO 2	Valor do campo dedicado ao armazenamento de dados não solicitados, mas que o fornecedor eventualmente disponha (ex. número de eixos)	≤ 255	Alfanumérico  Não pode conter o caractere ‘;’
<b>22</b>	NOME CAMPO RESERVADO 2	Nome do campo	≤ 255	Não pode conter o caractere ‘;’
<b>23</b>	CAMPO RESERVADO 3	Valor do campo dedicado ao armazenamento de dados não solicitados, mas que o fornecedor eventualmente disponha (ex. número de eixos)	≤ 255	Alfanumérico  Não pode conter o caractere ‘;’
<b>24</b>	NOME CAMPO RESERVADO 3	Nome do campo	≤ 255	Alfanumérico  Não pode conter o caractere ‘;’

<b>25</b>	<b>CAMPO RESERVADO 4</b>	Valor do campo dedicado ao armazenamento de dados não solicitados, mas que o fornecedor eventualmente disponha (ex. número de eixos)	≤ 255	Alfanumérico Não pode conter o caractere ‘;’
<b>26</b>	<b>NOME CAMPO RESERVADO 4</b>	Nome do campo	≤ 255	Alfanumérico Não pode conter o caractere ‘;’
<b>27</b>	<b>CAMPO RESERVADO 5</b>	Valor do campo dedicado ao armazenamento de dados não solicitados, mas que o fornecedor eventualmente disponha (ex. número de eixos)	≤ 255	Alfanumérico Não pode conter o caractere ‘;’
<b>28</b>	<b>NOME CAMPO RESERVADO 5</b>	Nome do campo	≤ 255	Alfanumérico Não pode conter o caractere ‘;’

### Observações importantes:

- Os campos devem ser separados por ‘;’, e este caractere NÃO PODE aparecer em nenhum outro campo, mesmo campos de texto, como de “Observações”.
- Os campos que o fornecedor não esteja apto a extrair (ou não se aplique), deixar o campo vazio (“”). Assim, supondo (exemplo de caráter ilustrativo, apenas) que haja 4 somente 4 campos, o primeiro tenha valor “1”, o segundo valor “1000”, o terceiro o operador do pórtico não esteja apto a extrair, o quarto seja igual a “99”, a linha a ser gravada no arquivo texto será:  
1;1000;;99  
Note que não há espaços entre os 2 ponto-e-vírgula do campo não disponível.
- Campos que o fornecedor possa extrair, mas tenha ocorrida falha, anotar o campo o valor “-1”
- Separador decimal será a vírgula “,”
- Não utilizar separador de milhar
- Não utilizar acentuação e/ou “;” nos campos alfanuméricos
- O comprimento do veículo é o melhor valor estimado e informado pelos sensores LAP ou laser.
- Os campos reservados destinam-se ao armazenamento de informações não solicitadas, mas que eventualmente o fornecedor disponha. Há cinco campos reservados que podem ser empregados para este fim.
- Na descrição do campo reservado, o fornecedor deve especificar o qu é o campo.

### Arquivo de Imagem

As imagens deverão ser enviadas em formato JPG com tamanho em bytes que exceda 300 Kbytes. O nome do arquivo de imagem terá o seguinte formato:

Formato:

<ID\_LOCAL>\_<AA><MM><DD><HH><MM><SS>\_<P>\_<SN>\_<SEQ\_NUM>.JPG

Exemplo: Imagem do pórtico de Barra Mansa, feita em 21 do mês 05 de 2012, as 01h53m55s na faixa 1, no equipamento com Número de série XYZ001, e foto de número sequencial 14322:

*BMA\_20120521015355\_1\_XYZ001\_0000014322.JPG*

Tabela 2a. Campos que devem conter no nome do arquivo de imagem.

#	Campo	Descrição	Número de caracteres	Tipo original
1	ID_LOCAL	Identificação do fornecedor	3	Alfabético
2	AAAA	Ano	4	Numérico inteiro
3	MM	Mês	2	Numérico inteiro
4	DD	Dia	2	Numérico inteiro
5	HH	Hora	2	Numérico inteiro
6	MM	Minutos	2	Numérico inteiro
7	SS	Segundos	2	Numérico inteiro
	F	Faixa de rodagem de tráfego (número da faixa, da esquerda para a direita)	1	Numérico inteiro
	SN_EQUIPTO	Identificação única (serial number) do equipamento no qual foi gerado o log. Considera-se um equipamento (câmera) por pista.	≤ 255	Alfabético
	SEQ_NUM	Número sequencial por faixa de rodagem, referente ao número da imagem disparada pela câmera.	7	Numérico longo inteiro ( <i>long int</i> )

## Identificação do fornecedor/local do pórtico

Cada pórtico sob responsabilidade do fornecedor terá um código único de três letras, representado pela sua localidade, conforme a tabela:

Tabela 3a. Siglas identificadoras dos locais de instalação dos pórticos (ID\_LOCAL).

ID_LOCAL	Localização
----------	-------------

GRU	Rodovia Dutra – Guarulhos KM 227
SJC	Rodovia Dutra – São José dos Campos KM 154
BMA	Barra Mansa – RJ KM 266
SJM	São José do Meriti – RJ KM 167

## Data e Horário

### 1.6.Intervalos de horário

A tabela mostra os intervalos de tempo a serem considerados no sistema.

Tabela 4a. Intervalos de tempo considerados para gravação de registros em arquivos.

<b>Intervalo</b>	<b>Horário Início</b>	<b>Horário fim</b>
<b>01</b>	00:00:00,000	00:29:59,999
<b>02</b>	00:30:00,000	00:59:59,999
<b>03</b>	01:00:00,000	01:29:59,999
<b>04</b>	01:30:00,000	01:59:59,999
<b>05</b>	02:00:00,000	02:29:59,999
<b>06</b>	02:30:00,000	02:59:59,999
<b>07</b>	03:00:00,000	03:29:59,999
<b>08</b>	03:30:00,000	03:59:59,999
<b>09</b>	04:00:00,000	04:29:59,999
<b>10</b>	04:30:00,000	04:59:59,999
<b>11</b>	05:00:00,000	05:29:59,999
<b>12</b>	05:30:00,000	05:59:59,999
<b>13</b>	06:00:00,000	06:29:59,999
<b>14</b>	06:30:00,000	06:59:59,999
<b>15</b>	07:00:00,000	07:29:59,999
<b>16</b>	07:30:00,000	07:59:59,999
<b>17</b>	08:00:00,000	08:29:59,999
<b>18</b>	08:30:00,000	08:59:59,999
<b>19</b>	09:00:00,000	09:29:59,999

<b>Intervalo</b>	<b>Horário Início</b>	<b>Horário fim</b>
<b>20</b>	09:30:00,000	09:59:59,999
<b>21</b>	10:00:00,000	10:29:59,999
<b>22</b>	10:30:00,000	10:59:59,999
<b>23</b>	11:00:00,000	11:29:59,999
<b>24</b>	11:30:00,000	11:59:59,999
<b>25</b>	12:00:00,000	12:29:59,999
<b>26</b>	12:30:00,000	12:59:59,999
<b>27</b>	13:00:00,000	13:29:59,999
<b>28</b>	13:30:00,000	13:59:59,999
<b>29</b>	14:00:00,000	14:29:59,999
<b>30</b>	14:30:00,000	14:59:59,999
<b>31</b>	15:00:00,000	15:29:59,999
<b>32</b>	15:30:00,000	15:59:59,999
<b>33</b>	16:00:00,000	16:29:59,999
<b>34</b>	16:30:00,000	16:59:59,999
<b>35</b>	17:00:00,000	17:29:59,999
<b>36</b>	17:30:00,000	17:59:59,999
<b>37</b>	18:00:00,000	18:29:59,999
<b>38</b>	18:30:00,000	18:59:59,999
<b>39</b>	19:00:00,000	19:29:59,999
<b>40</b>	19:30:00,000	19:59:59,999
<b>41</b>	20:00:00,000	20:29:59,999
<b>42</b>	20:30:00,000	20:59:59,999
<b>43</b>	21:00:00,000	21:29:59,999
<b>44</b>	21:30:00,000	21:59:59,999
<b>45</b>	22:00:00,000	22:29:59,999
<b>46</b>	22:30:00,000	22:59:59,999
<b>47</b>	23:00:00,000	23:29:59,999
<b>48</b>	23:30:00,000	23:59:59,999

## Registro de LOG

Para fins de levantamento de dados, será considerado o registro de eventos do sistema contendo o horário, o equipamento (identificado por pista) e a descrição do erro ocorrido, ocorrência ou aviso do log respectivo. Serão consideradas ocorrências os eventos de desligamento e ligação do equipamento, reinício forçado (watchdog), erros ou avisos pertinentes no sistema.

Formato:

*<ID\_LOCAL>;<AAAA>;<MM>;<DD>;<HH>;<MM>;<SS>;<mmmm>;<SN\_EQUIPTO>;<P>;<LOG>;*

Exemplo:

*GRU;2012;05;20;13;41;32;5221;SN1101A;2; O sistema foi reiniciado manualmente pelo operador Mario.*

Tabela 5a. Tabela dos campos de LOG.

#	Campo	Descrição	Número de caracteres	Tipo original
1	ID_LOCAL	Identificação do fornecedor	3	Alfabético
2	AAAA	Ano	4	Numérico inteiro
3	MM	Mês	2	Numérico inteiro
4	DD	Dia	2	Numérico inteiro
5	HH	Hora	2	Numérico inteiro
6	MM	Minutos	2	Numérico inteiro
7	SS	Segundos	2	Numérico inteiro
8	mmm	Milissegundos	3	Numérico inteiro
9	SN_EQUIPTO	Identificação única (serial number) do equipamento no qual foi gerado o log. Considera-se um equipamento (câmera) por pista.	≤ 255	Alfabético
10	F	Faixa de rodagem de tráfego (número da faixa, da esquerda para a direita)	1	Numérico inteiro
11	LOG	Descrição do erro ocorrido, ocorrência ou aviso do LOG respectivo	≤ 255	Alfanumérico



Os arquivos de LOG serão diários e formato de arquivo texto. O arquivo ficará localizado dentro do Diretório Diário. “A nomenclatura seguirá a mesma do diretório principal, com a extensão de arquivo “.LOG”.

Formato:

<ID\_LOCAL>\_<AAAA>\_<MM>\_<DD>.LOG

Exemplo:

BMA\_2012\_05\_20.LOG

## **Apêndice 2 – Exemplo numérico de cálculo de índices e frações na estatística de inspeção de imagens.**

### **Funções de dependência entre notas %ICE, %ICH e novas definições.**

Definem-se dois índices %IPI e %ISP:

☞ %IPI = índice de imagens com placas ilegíveis (imagem contém uma placa que é ilegível) do total de placas não lidas;

☞ %IPP = índice de imagens sem placa (nenhuma placa pode ser reconhecida na imagem<sup>37</sup>) do total de placas não lidas.

Sem exibir a dependência com os subscritos j e k, não é difícil ver que as seguintes relações são válidas (onde os índices %fCH, %fPI e %fII foram definidos na Seção “Hierarquia binomial usada e notas finais”):

$$\%fCE = \%ICE,$$

$$\%fCH = \%ICH(1 - \%ICE),$$

$$\%fPI = \%IPI(1 - \%ICE)(1 - \%ICH),$$

$$\%fPP = (1 - \%IPI)(1 - \%ICE)(1 - \%ICH).$$

$$\%fII = \%fPI + \%fPP,$$

Ou seja, a partir da estimativa dos índices é possível calcular as frações. É importante ver que as frações assim calculadas permitem que se determinem os intervalos de erro a partir dos desvios padrão para os índices que estão bem definidos por conta da estatística de distribuição binomial. Uma vez determinados os desvios padrão de cada índice  $\sigma_{ICE}$ ,  $\sigma_{ICH}$ ,  $\sigma_{IPI}$  etc, então podemos determinar os erros dos índices:

---

<sup>37</sup> Isso se deve a problemas diversos: iluminação, problema de trigger, posicionamento do veículo etc.

$$\sigma_{fCE} = \sigma_{ICE},$$

$$\sigma_{fCH} = (1 - \%ICE)\sigma_{ICH} + \%ICH\sigma_{ICE},$$

$$\sigma_{fPI} = (1 - \%ICE)(1 - \%ICH)\sigma_{IPI} + \%IPI(1 - \%ICH)\sigma_{ICE} + (1 - \%ICE)(1 - \%IPI)\sigma_{ICH},$$

$$\sigma_{fPP} = \sigma_{fPI},$$

A sequência de planilhas abaixo feitas em Mathcad (Wikipedia, 2012) simulam a estatística para as notas de perda definidas aqui e as notas principais %ICE e %ICH.

Distribuição Uniforme

Vetor de variáveis distribuídas uniformemente entre 0 e 1

$n := 120$       Número de elementos do vetor       $i := 0..n-1$

$x_0 := 0$        $x_1 := 1$

$U := \text{runif}(n, 0, 1)$       Esta subrotina cria números aleatórios com distribuição uniforme entre 0 e 1.

$R(i, t) := \text{if}(U_i \geq t, 0, 1)$       Esta função cria uma sequência de "0" e "1" conforme a estatística da taxa t.

Taxa de acerto       $t := 65\%$       A taxa aqui é um parâmetro dado. O objetivo é determinar a taxa simulada.

$r_i := R(i, t)$       Vetor com a sequência gerada

Para verificar o resultado, simplesmente contamos todos os '1' da amostra

$$\text{soma\_acertos} := \sum_{k=0}^{n-1} r_k$$

$\text{soma\_acertos} = 77.0000$

Portanto a taxa 'simulada' de acerto será       $t_{\text{exp}} := \frac{\text{soma\_acertos}}{n}$        $t_{\text{exp}} = 64.17\%$

Figura 79 Sequência de código do script em MathCAD para simular uma sequência com distribuição binomial para a taxa simulada t. Neste exemplo, t=65%, enquanto que a taxa simulada (razão entre a contagem de '1's e o número total da amostra) é 64,17%.

### Subrotina para determinação dos índices e frações

A sub-rotina da Figura 80 e Figura 81 mostra a sequência de comandos gerados para simular os índices ICE, ICH, IPI e IPP, bem como os desvios-padrão. Uma vez determinados os índices, as frações (ver “Hierarquia binomial usada e notas finais”) podem ser facilmente calculadas por meio da sub-rotina da Figura 82.

```

SimulIndices(n, t, u, v) :=
  P ← runif(n, 0, 1)
  for i ∈ 0..n-1
    pi ← if(Pi ≥ t, 0, 1)
  nICE ←  $\sum_{i=0}^{n-1} p_i$ 
  t%ICE ←  $\frac{nICE}{n}$ 
  Q ← runif(n - nICE, 0, 1)
  for i ∈ 0..n - nICE - 1
    qi ← if(Qi ≥ u, 0, 1)
  nICH ←  $\sum_{i=0}^{n-nICE-1} q_i$ 
  t%ICH ←  $\frac{nICH}{n - nICE}$ 
  R ← runif(n - nICE - nICH, 0, 1)
  for i ∈ 0..n - nICE - nICH - 1
    ri ← if(Ri ≥ v, 0, 1)
  nIPI ←  $\sum_{i=0}^{n-nICE-nICH-1} r_i$ 
  t%IPI ←  $\frac{nIPI}{n - nICE - nICH}$ 
  t%IPP ← 1 - t%IPI
  X0 ← t%ICE
  X1 ← t%ICH
  X2 ← t%IPI
  X3 ← t%IPP

```

Figura 8o Sub-rotina para geração dos índices ICE, ICH, IPI e IPP a partir de n elementos em uma amostra com p = t, q = u e r = v fornecidas no cabeçalho da função.

$$\begin{aligned}
 X_4 &\leftarrow \sqrt{\frac{t_{\%ICE} \cdot (1 - t_{\%ICE})}{n}} \\
 X_5 &\leftarrow \sqrt{\frac{t_{\%ICH} \cdot (1 - t_{\%ICH})}{n - n_{ICE}}} \\
 X_6 &\leftarrow \sqrt{\frac{t_{\%IPI} \cdot (1 - t_{\%IPI})}{n - n_{ICE} - n_{ICH}}} \\
 X_7 &\leftarrow X_6 \\
 X &
 \end{aligned}$$

Figura 81 Cálculo dos desvios-padrão para a saída da função SimulIndices(n,t,u,v).

```

SimulFracoes(n,t,u,v) :=
  X ← SimulIndices(n,t,u,v)
  F0 ← X0
  F1 ← X1 · (1 - X0)
  F2 ← X2 · (1 - X1) · (1 - X0)
  F3 ← (1 - X2) · [(1 - X1) · (1 - X0)]
  F4 ← X4
  F5 ← (1 - X0) · X5 + X1 · X4
  F6 ← X6 · (1 - X1) · (1 - X0) + X2 · (1 - X0) · X5 + X2 · (1 - X1) · X4
  F7 ← X6 · (1 - X1) · (1 - X0) + (1 - X2) · (1 - X0) · X5 + (1 - X2) · (1 - X1) · X4
  G(0) ← X
  G(1) ← F
  G

```

Figura 82 Sub-rotina para determinação das frações e seus desvios-padrão.

## Simulações numéricas

Se admitirmos que  $t = 50\%$ ,  $u = 50\%$  e  $v = 50\%$ , então esperamos encontrar, do total de  $n$  imagens, que  $p = 50\%$  aproximadamente é formado por imagens que foram lidas corretamente pelo software de OCR, que  $q = 50\%$  das imagens, do total de imagens não lidas corretamente pelo software OCR, foram inspecionadas corretamente pelos humanos e que,  $r = 50\%$  das imagens, do total que não foi lido corretamente pelos humanos, contém placas ilegíveis. Portanto  $s = 50\%$  desse total contém imagens sem placas. Não é difícil ver que, *do total de imagens*, 50% foi lida corretamente, 25% pode apenas ser lida por humanos, 12,5% é formado por imagens que contém placa, mas não podem ser lidas nem por humanos e, finalmente, que 12,5% contém imagens sem placas.

Ao se rodar uma instância da sub-rotina Simulfracoes() para  $n = 120$  elementos com essas taxas encontramos:

$$\begin{pmatrix} 47.5 & 47.5 \\ 60.3 & 31.7 \\ 56.0 & 11.7 \\ 44.0 & 9.2 \\ 4.6 & 4.6 \\ 6.2 & 6.0 \\ 9.9 & 4.9 \\ 9.9 & 4.3 \end{pmatrix} \%$$

ou seja,

$$p=47.5\pm4.6\%$$

$$q=60.3\pm6.2\%$$

$$r=56.0\pm9.9\%$$

$$s=44.0\pm9.9\%$$

Os índices e desvios correspondentes são:

$$\%fCE = 47.5\pm4.6\%$$

$$\%fCH = 31.7\pm6.0\%$$

$$\%fPI=11.7\pm4.9\%$$

$$\%fPP=9.2\pm4.3\%$$

Se usarmos  $n=1200$  elementos, o resultado mostra uma redução grande nos desvios, com aumento da acurácia no cálculo dos índices e frações:

$$\begin{pmatrix} 46.8 & 46.8 \\ 50.5 & 26.9 \\ 49.4 & 13.0 \\ 50.6 & 13.3 \\ 1.4 & 1.4 \\ 2.0 & 1.8 \\ 2.8 & 1.6 \\ 2.8 & 1.6 \end{pmatrix} \%$$

ou seja,

$$p=46.8\pm1.4\%$$

$$q=50.5\pm2.0\%$$

$$r=49.4\pm2.8\%$$

$$s=50.6\pm2.8\%$$

Os índices e desvios correspondentes são:

$$\%fCE = 46.8\pm1.4\%$$

$$\%fCH = 26.9\pm1.8\%$$

$$\%fPI=13.0\pm1.6\%$$

$$\%fPP=13.3\pm1.6\%$$

Esses valores devem ser comparados as valores fornecidos como parâmetros:  $p = 50\%$ ,  $q = 50\%$ ,  $r = 50\%$ ,  $s = 50\%$ ,  $\%fCE = 50\%$ ,  $\%fCH=25\%$ ,  $\%fPI = 12,5\%$  e  $\%fPP=12,5\%$ .

Na simulação numérica, é demonstrado oscilações das estimativas das frações  $\%fCE$ ,  $\%fCH$  e  $\%fPP$  como função do número de elementos da amostra na Figura 83.

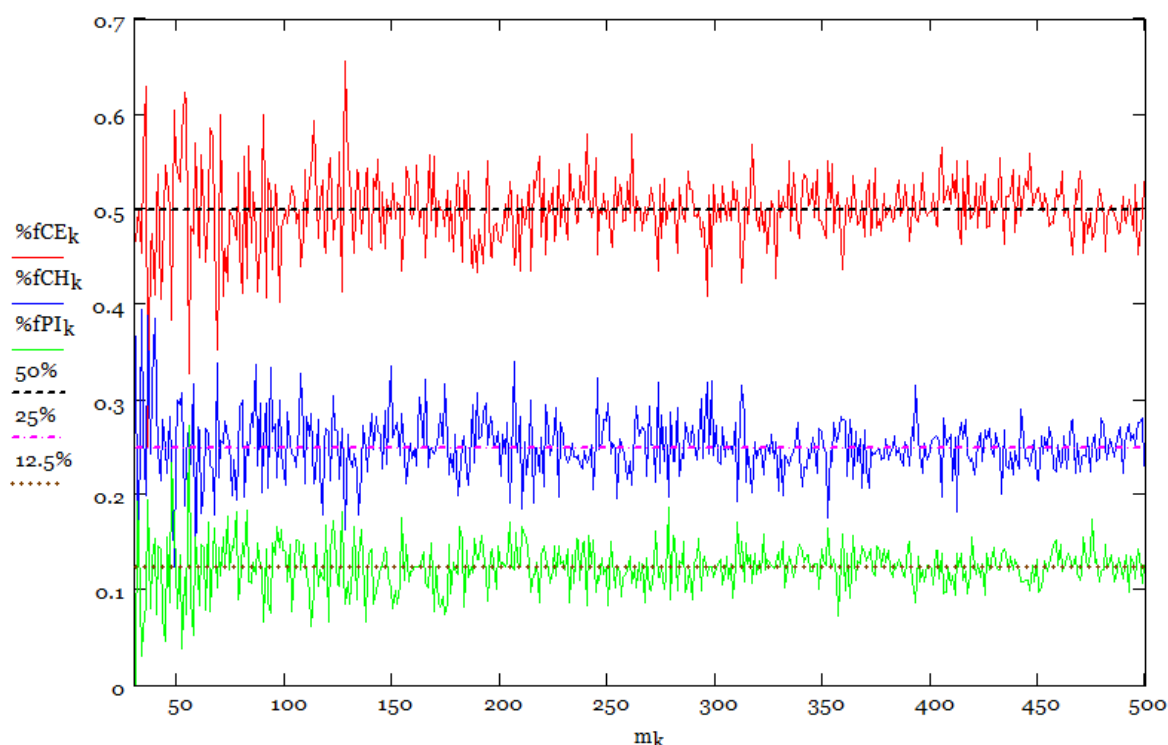


Figura 83 Oscilação das estimativas das frações como função do número de elementos da amostra.

A dependência dos desvios padrão com o número de elementos da amostra, tanto para os índices como para as frações, pode ser vista na Figura 84 e Figura 85.

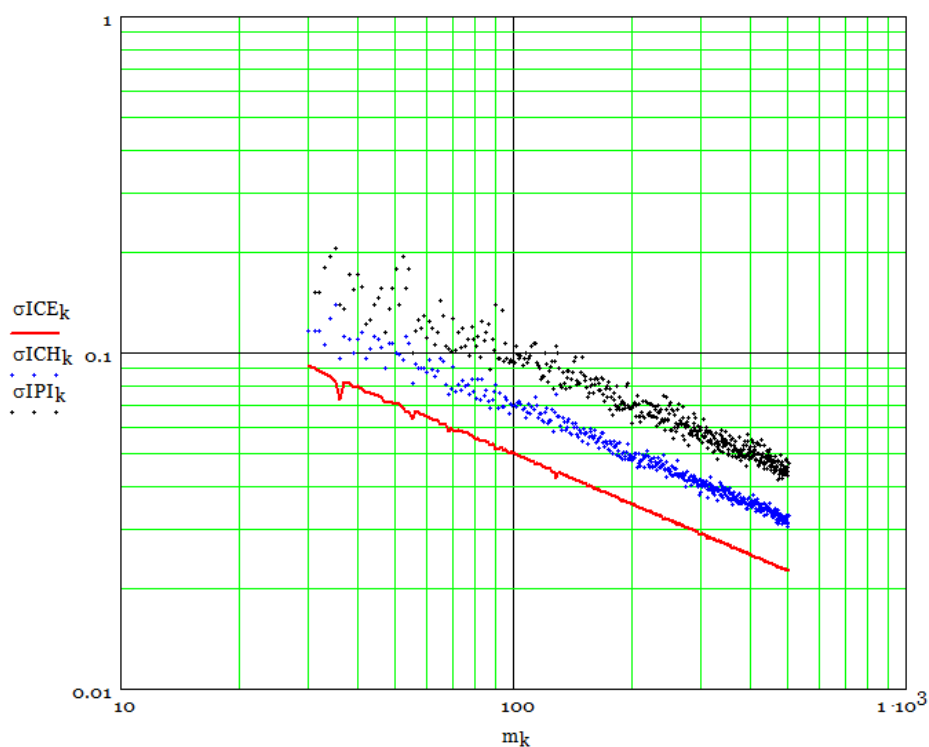


Figura 84 Dependência dos desvios padrão para os 3 índices principais (ICE, ICH e IPI – o desvio para IPP é igual a IPI) como função do número de elementos da amostra (simulação).

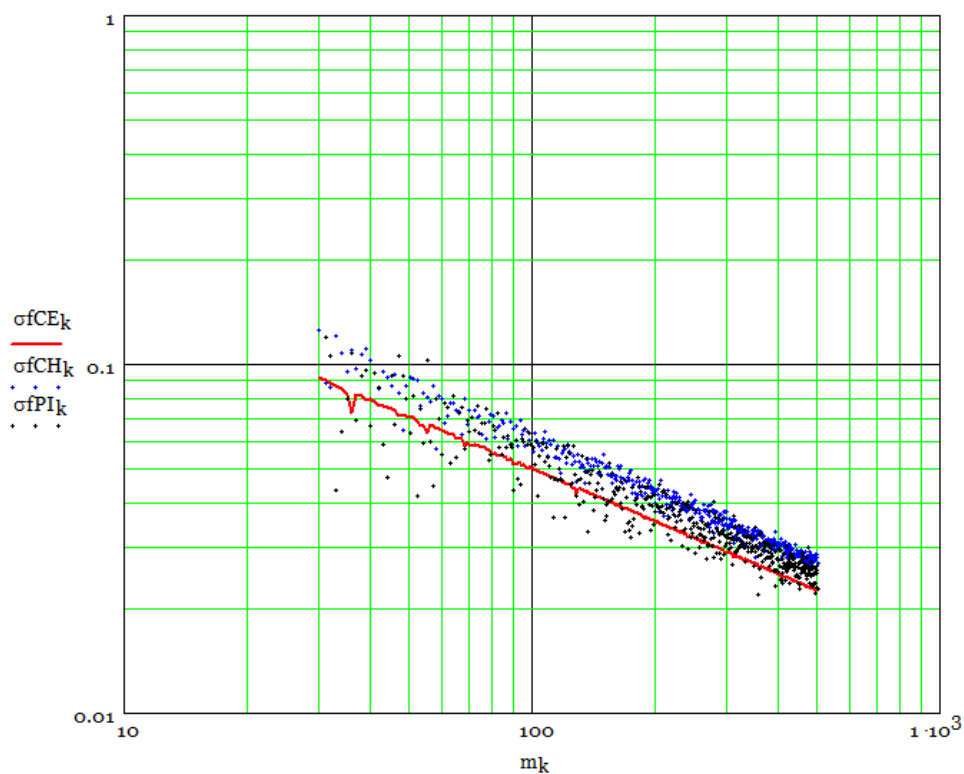


Figura 85 Dependência dos desvios padrão para as 3 frações principais (fCE, fCH e fPI – o desvio para fPP é igual a fPI) como função do número de elementos da amostra (simulação).



## **Apêndice 3 – Analise restrita de fluxo, velocidades instantâneas e densidade de tráfego para GRU.**

No que segue, apresentamos alguns histogramas de distribuição de velocidade e comprimento para o sítio GRU que já constam nos relatórios de registro de passagem. Não realizaremos a análise da dependência das notas com a densidade<sup>38</sup> porque esse estudo não está disponível para todos os sítios, o que limita a variabilidade em geometria e posicionamento dos pontos. O que mostramos a seguir é apenas um registro restrito do que seria possível fazer em termos de análise de dados, caso todos os fornecedores tivessem os mesmos recursos instalados.

---

<sup>38</sup> Mesmo que as frações sejam acuradamente determinadas, variações extrínsecas como a densidade (velocidade) podem causar modificações de desempenho que se refletem nessas frações.

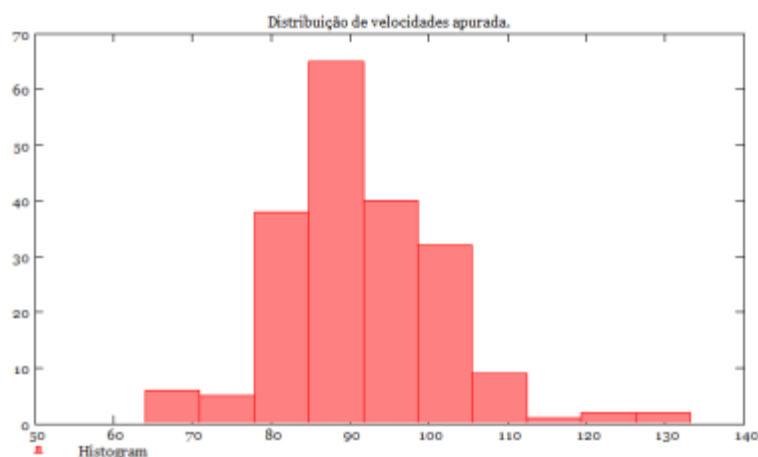


Figura 86 Distribuição de velocidades instantâneas (em km/h) dos veículos em GRU em 1/9/2012 das 9:00 as 9:21.

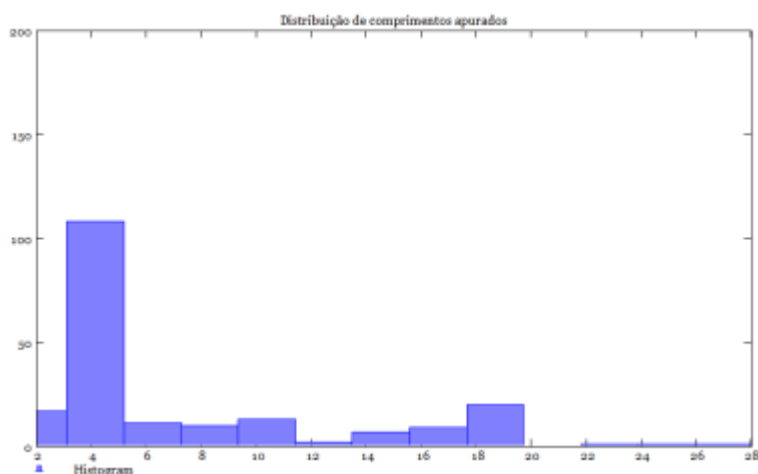


Figura 87 Distribuição de comprimentos de veículos (em metros) em GRU em 1/9/2012 das 9:00 as 9:21. Diversas classes podem ser observadas nesta distribuição.

A Figura 87 ilustra um gráfico da contagem dos veículos por velocidade em um período matutino em GRU em 1/9/2012 que detectou velocidades mínimas da ordem de 65km/h e máximas da ordem de 130km/h. A velocidade mais provável praticada no trecho é 90km/h enquanto que a média das velocidades (centroide da distribuição) é 91,11km/h. A Figura 88 ilustra a distribuição de comprimentos em metros. A maioria dos veículos tem até 4 metros de comprimento (o que corresponde aos carros). É possível ver uma classe que se distingue: a de veículos com dimensão superior a 22 metros. Para que tenhamos uma ideia de como o comprimento correlaciona com a velocidade, a Figura 88 traz um ‘scatter plot’ correspondente. Os diversos períodos de tarde e noite são apresentados a seguir.

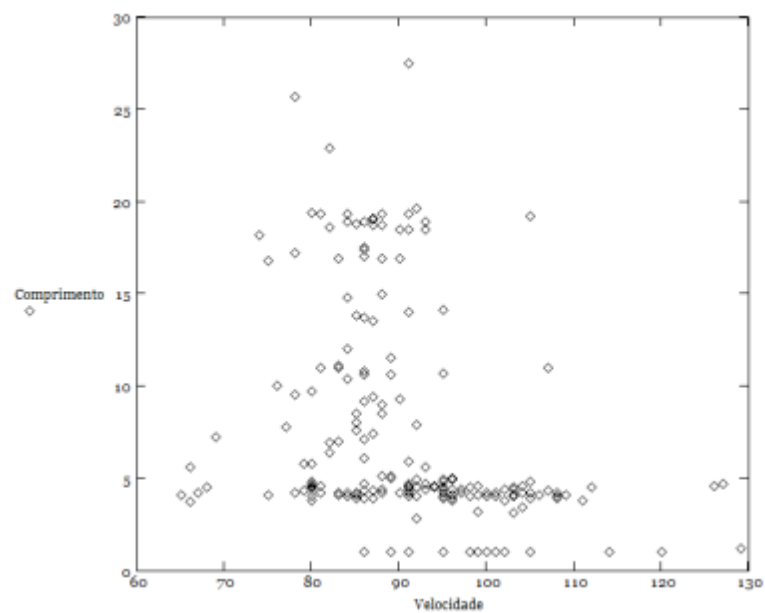


Figura 88 Gráfico de pontos que mostram a distribuição Comprimento [metros] X velocidade [km/h] apurada no período de 9:00 a 9:21 em 1/9/2012 em GRU. A sequência de pontos na parte inferior corresponde aos veículos do tipo ‘moto’. A sequência acumulada em torno de 4 metros é o grupo ‘carro’.

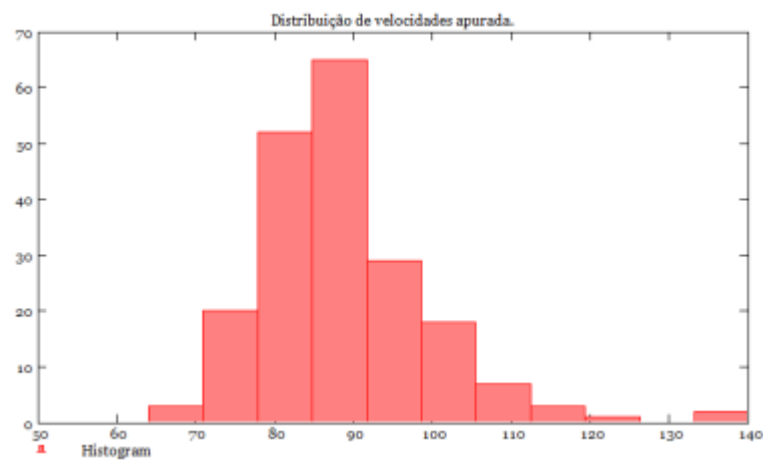


Figura 89 Distribuição de velocidades instantâneas (em km/h) dos veículos em GRU em 1/9/2012 das 14:00 as 14:16. A velocidade média foi de 88,43km/h durante esse período.

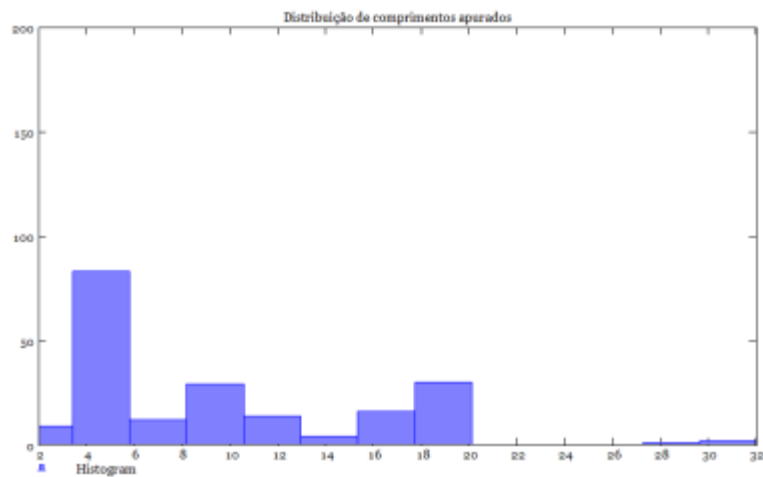


Figura 90 Distribuição de comprimentos de veículos (em metros) em GRU em 1/9/2012 das 14:00 as 14:16. Diversas classes podem ser observadas nesta distribuição.

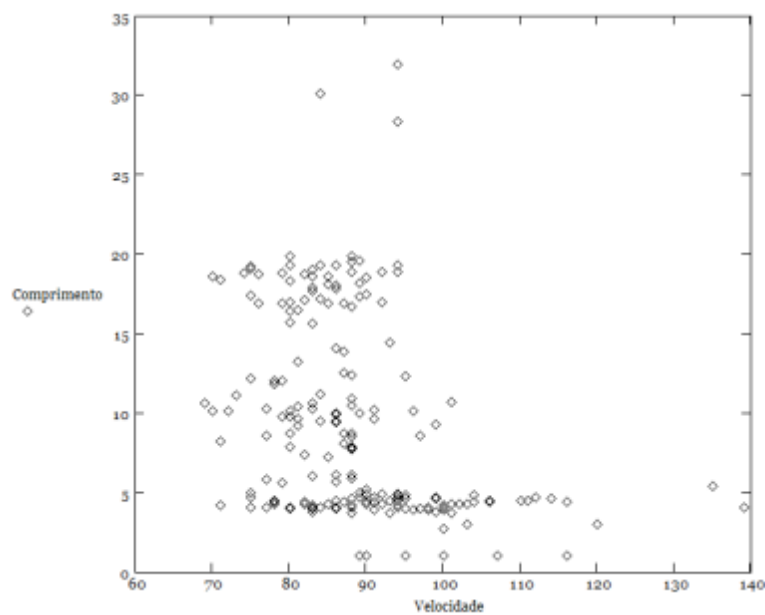


Figura 91 Gráfico de pontos que mostram a distribuição Comprimento [metros] X velocidade [km/h] apurada no período de 14:00 a 14:16 em 1/9/2012 em GRU. A sequência de pontos na parte inferior corresponde aos veículos do tipo 'moto'. A sequência acumulada em torno de 4 metros é o grupo 'carro'.

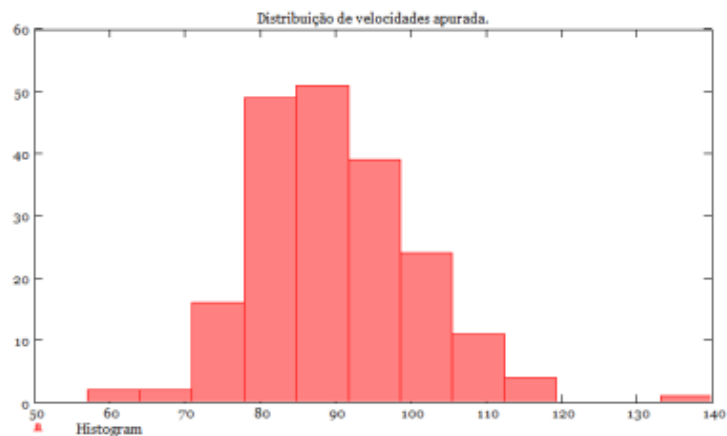


Figura 92 Distribuição de velocidades instantâneas (em km/h) dos veículos em GRU em 1/9/2012 das 19:00 as 19:19. A velocidade média do ponto foi de 89,89km/h durante esse período.

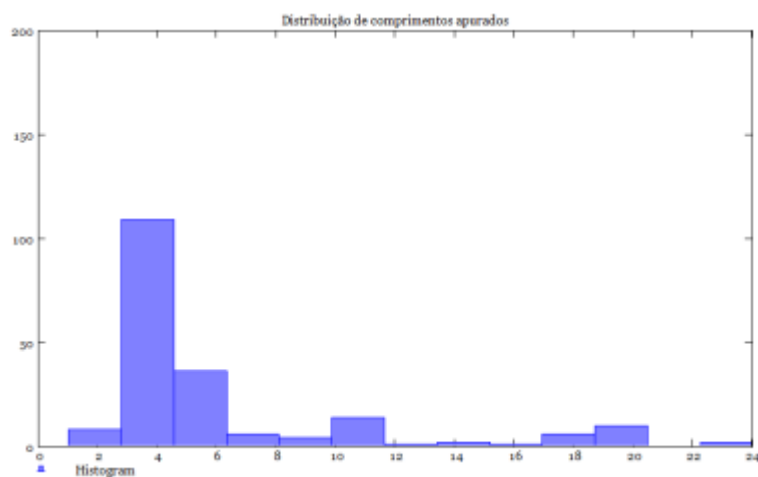


Figura 93 Distribuição de comprimentos de veículos (em metros) em GRU em 1/9/2012 das 19:00 as 19:19. Diversas classes podem ser observadas nesta distribuição.

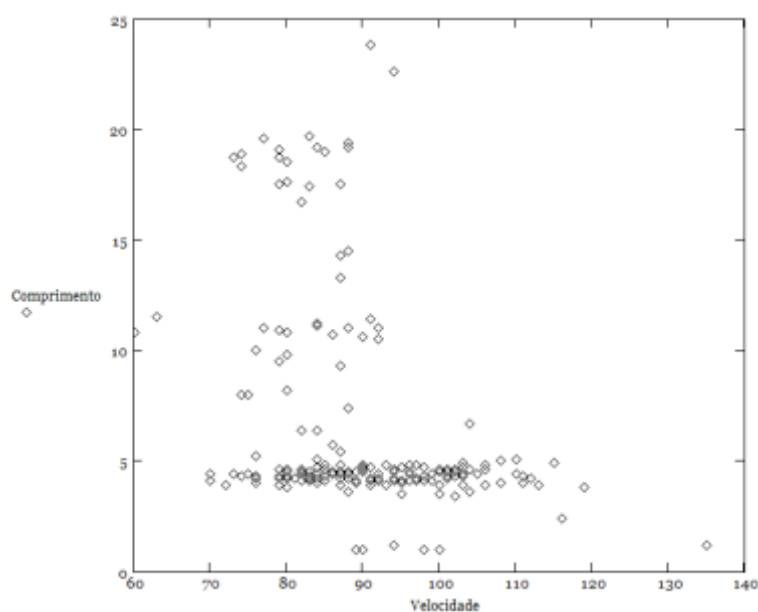


Figura 94 Gráfico de pontos que mostram a distribuição Comprimento [metros] X velocidade [km/h] apurada no período de 19:00 a 19:19 em 1/9/2012 em GRU. A sequência de pontos na parte inferior

corresponde aos veículos do tipo ‘moto’. A sequência acumulada em torno de 4 metros é o grupo ‘carro’.

Outra análise que foi possível com os dados do pórtico de Guarulhos foi o de velocidade *por pista*. A Figura 95 ilustra uma imagem tirada no dia 28/3/2012 do tráfego sobre a passarela onde foram instalados os equipamentos do ponto de Guarulhos.



Figura 95 Foto em Março de 2012 (antes do início do projeto) do tráfego sobre a passarela no pórtico do km 227 de Guarulhos.

A numeração das pistas segundo essa imagem é da direita para a esquerda, sendo que as pistas 1 e 2 estão na chamada ‘via expressa’ da Dutra (ver Figura 95). As velocidades médias inferidas no dia 6/9/2012 em períodos selecionados da tarde e manhã para cada uma das pistas nessa numeração podem ser vistos na Figura 97 e Figura 98.

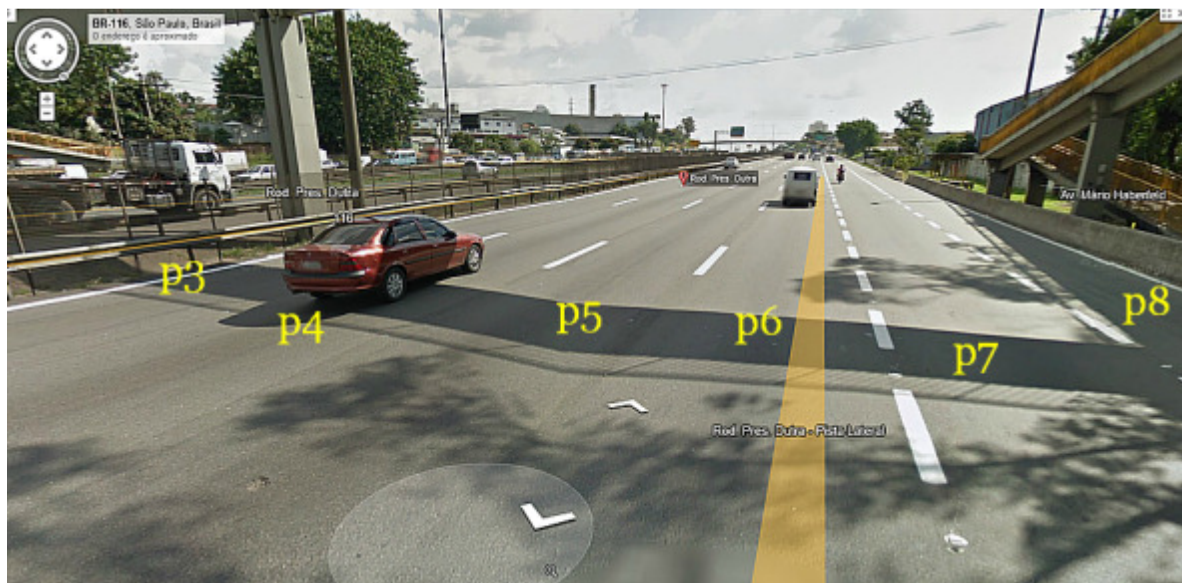


Figura 96 Identificação das pistas sob o ponto de aferição na marginal sentido Sul da rodovia Pres. Dutra no km 227 de Guarulhos. A pista 8 é a do canto à esquerda e seu volume de tráfego é muito reduzido. Fonte: Google.

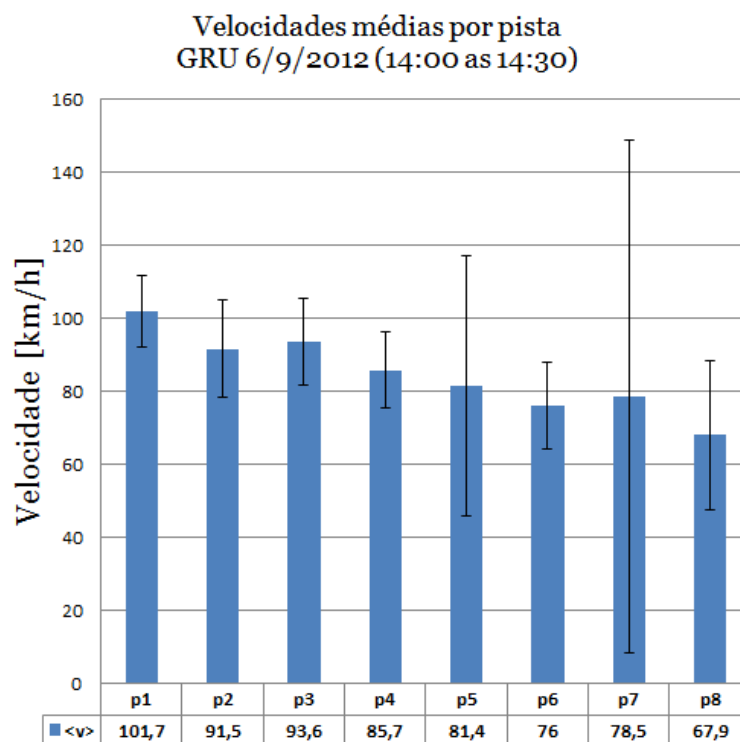


Figura 97 Velocidades médias e desvios-padrão no pórtico de Guarulhos como função da pista em um período vespertino indicado.



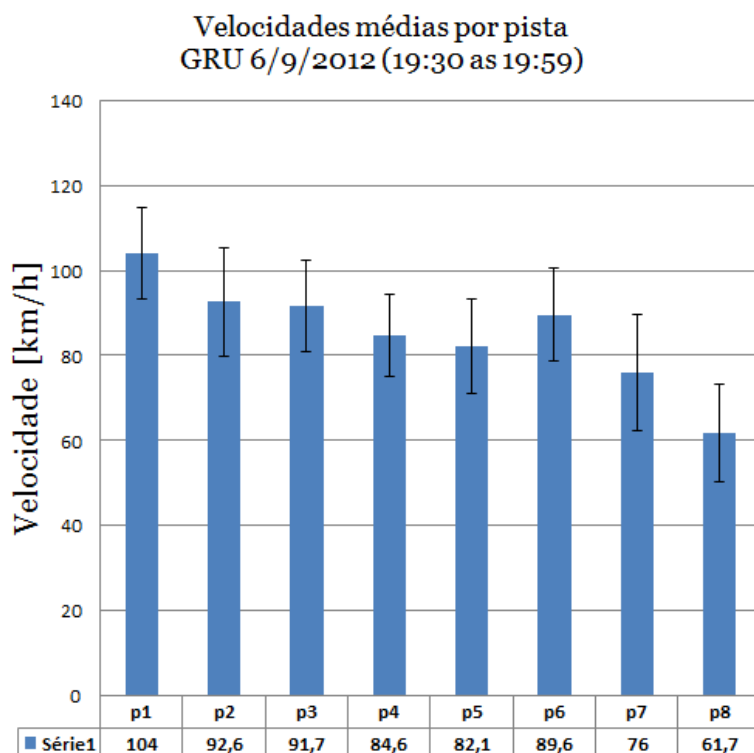


Figura 98 Velocidades médias e desvios-padrão no pórtico de Guarulhos como função da pista em um período noturno indicado.

O sistema instalado capturou uma deriva de até 40% na velocidade em direção às pistas de maior numeração (mais próximas do acostamento). É importante considerar que a estatística mostrada não foi feita com o mesmo número de veículos. Para os períodos escolhidos, a Figura 99 mostra os fluxos em cada pista (tarde/noite nos períodos considerados), o que confere as pistas 3, 4 e 5 os maiores volumes (aproximadamente 17,4% do volume total passa por essas pistas, o que representa cerca de 50% do total do volume). Menos de 0,4% do volume total passa pela pista 8.

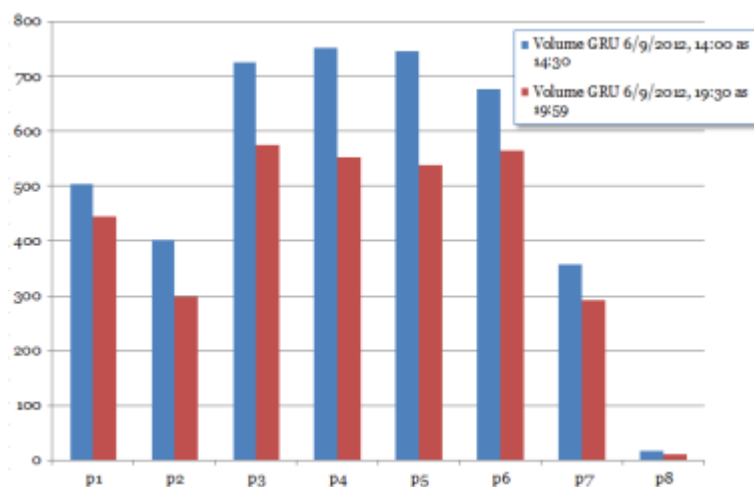


Figura 99 Volumes (fluxos) de tráfego por pista conforme o período considerado na legenda.

A ocupação de veículos longos (aqui definidos como veículos acima de 10 metros) é mostrada na Figura 100. Observamos que as pistas laterais das marginais

(e a pista 2) são as mais ocupadas na rodovia, o que mostra que a distribuição de retornos de um eventual sistema de pedágio baseado em OCR deve prover maior exatidão com as pistas “mais lentas”. Esse resultado mostra que a pista 8, apesar do baixo fluxo por se comportar como um ‘acostamento’, também deve ser monitorada. Isso acontece exatamente nesse local por causa do entroncamento das rodovias Presidente Dutra e Fernão Dias que continua, no acesso a Dutra, na forma da pista 8.

O perfil de excesso de velocidade (acima de 100km/h) é diferente entre os dois períodos considerados, conforme mostra a Figura 102. Em percentuais, a ocupação de veículos com excesso de velocidade é invertida da ocupação total ou volume de tráfego (Figura 99).

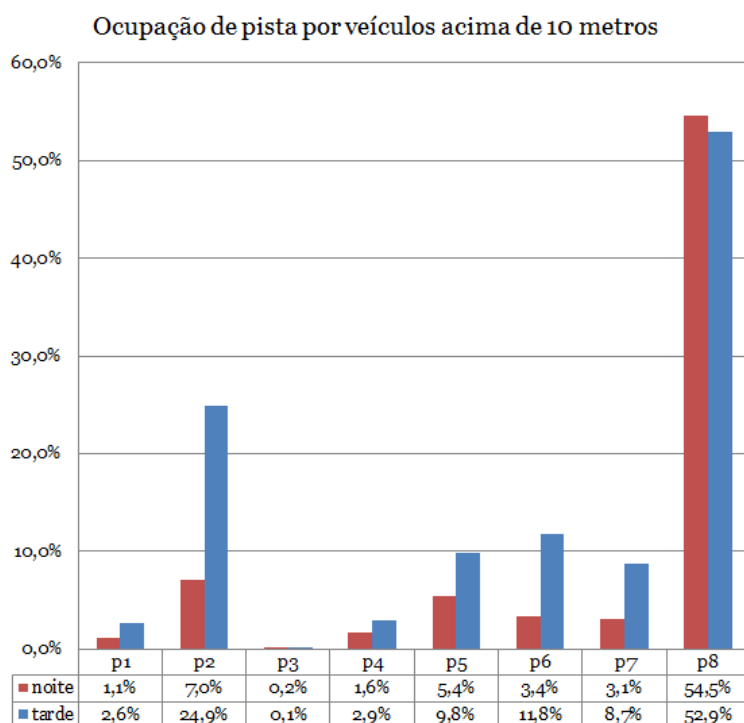


Figura 100 Distribuição de ocupação das pistas com veículos longos (acima de 10 metros) para os períodos de tarde e noite (das 14:00 as 14:30 e das 19:30 as 19:59 do dia 6/9/2012).

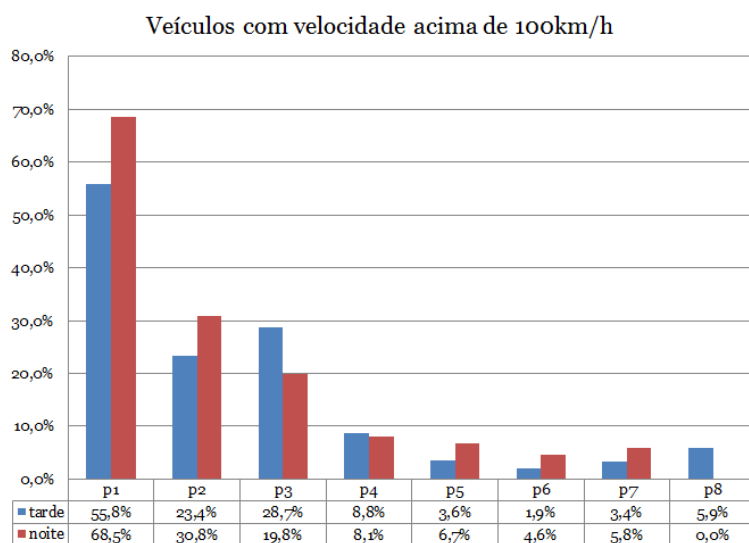


Figura 101 Distribuição de veículos com excesso de velocidade (acima de 100 km/h) para os períodos de tarde e noite (das 14:00 as 14:30 e das 19:30 as 19:59 do dia 6/9/2012).

Finalmente, podemos determinar as densidades. Para isso utilizamos os dados mostrados na tabela 4.

TARDE	vel med (km/h)	std vel	fluxo (veículos em 1/2 hora)	veículos/hora	Densidade (veículos/km)
p1	101,7	9,7	504	1008	9,9
p2	91,5	13,4	402	804	8,8
p3	93,6	11,9	726	1452	15,5
p4	85,7	10,4	752	1504	17,5
p5	81,4	35,5	746	1492	18,3
p6	76	12	676	1352	17,8
p7	78,5	70,2	357	714	9,1
p8	67,9	20,4	17	34	0,5
NOITE	vel med (km/h)	std vel	fluxo (veículos em 1/2 hora)	veículos/hora	Densidade (veículos/km)
p1	104	10,7	445	890	8,6
p2	92,6	12,7	299	598	6,5
p3	91,7	10,8	575	1150	12,5
p4	84,6	9,6	553	1106	13,1

p5	82,1	11,2	539	1078	13,1
p6	89,6	10,9	565	1130	12,6
p7	76	13,7	291	582	7,7
p8	61,7	11,5	11	22	0,4

Tabela 4 Velocidades médias, desvio padrão, fluxo (volumes em  $\frac{1}{2}$  hora), volume horário e a densidade de veículos (1/km) para o dia 6/9/2012 no ponto de Guarulhos.

A densidade pode ser calculada dividindo-se o fluxo horário pela velocidade média observada no período. Um histograma dessa densidade é visto na Figura 102.

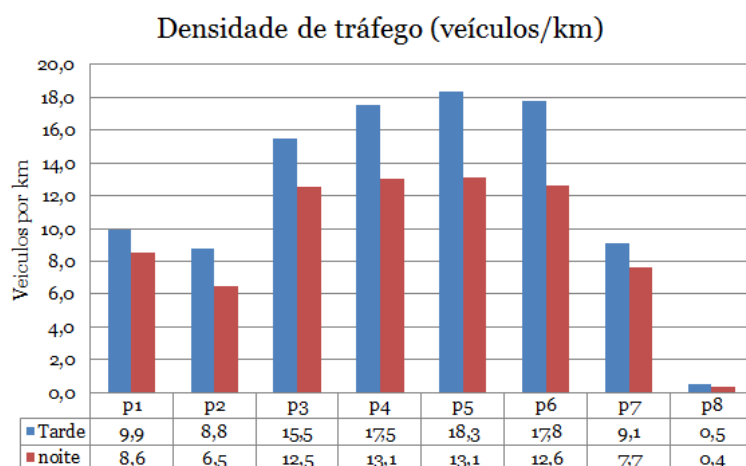


Figura 102 Distribuição de densidades de veículos (por km) no ponto de Guarulhos em 6/9 para os períodos da tarde (14:00 as 14:30) e noite (19:30 as 19:59).

Comentamos que, mesmo considerando veículos longos (com 30 metros de comprimentos) a densidade de tráfego observada máxima foi de 20 veículos por km, o que garante bastante espaçamento entre veículos de forma a não correr problemas de ocultamento. Como esperado, as densidades mostram que as pistas marginais (e dessas, as mais laterais) são as de maior densidade.

## Bibliografia

*Evaluation of the license plate recognition system.* (2012). Retrieved Agosto 2012, from [ssom.transportation.org: http://ssom.transportation.org/Documents/LPR\\_report\\_part3.pdf](http://ssom.transportation.org/Documents/LPR_report_part3.pdf)

(2012). *Sistema Automático de Arrecadação para Teste*. Jundiaí: Apresentação Engelog-CCR.

DENATRAN. (2010). *Resolução número 370 de 10 de dezembro de 2010*. Retrieved 8 2012, from [http://www.denatran.gov.br: http://www.denatran.gov.br/download/Resolucoes/RESOLUCAO\\_CONTRA N\\_370\\_10.pdf](http://www.denatran.gov.br/download/Resolucoes/RESOLUCAO_CONTRA_N_370_10.pdf)

DENATRAN. (2011). *Deliberação número 116 de 18 de outubro de 2011*. Retrieved 8 2012, from <http://www.denatran.gov.br>

- [http://www.denatran.gov.br/download/Deliberacoes/DELIBERACAO\\_CONT\\_RAN\\_116\\_11.pdf](http://www.denatran.gov.br/download/Deliberacoes/DELIBERACAO_CONT_RAN_116_11.pdf)
- DENATRAN/CGIE. (2006). *Manual de Procedimentos RENAVALAM*. Retrieved Agosto 2012, from <http://www.adotesc.com.br:>  
[http://www.adotesc.com.br/bd.arquivos/20100505manual\\_de\\_procediment.pdf](http://www.adotesc.com.br/bd.arquivos/20100505manual_de_procediment.pdf)
- Draghici, S. (1997). *A neural network based artificial vision system for licence plate recognition*. Retrieved Agosto 2012, from [vortex.cs.wayne.edu:](http://vortex.cs.wayne.edu:)  
<http://vortex.cs.wayne.edu/papers/ijns1997.pdf>
- Gazis, D. (2002). *Traffic Theory*. Kluwer Academic Publishers.
- Grant, J. (2012?). *Automatic License Plate Recognition*. Retrieved 8 15, 2012, from [www.nd.edu/~jgrant3/cw/alpr.pdf](http://www.nd.edu/~jgrant3/cw/alpr.pdf)
- Hadjipanayis, I. (2009). *Automatic License Plate Recognition (Project Report)*. University of Manchester .
- N. Sebe, I. C. (2005). *Machine Learning in Computer Vistion*. Berlin: Springer.
- Ozbay, S., & Ercelebi, E. (2005). Automatic Vehicle Identification by Plate Recognition. *World Academy of Science, Engineering and Technology* 9.
- Rossetti, M., & Baker, J. (2012?). *Applications and Evaluation of Automated License Plate Reading Systems*. Fayetteville: Dep. of Industrial Engineering, Univ. Arkansas.
- Silva, J., Marques, G., Jorge, P., Abrantes, A., Osório, A., Gomes, J., et al. (2006). Evaluation of an LPR-Based Toll Enforcement System on Portuguese Motorways. *Intelligent Transportation Systems Conference, ITSC 2006*, 719-724.
- Spiegel, M. (1975). *Probability and Statistics*. New York: McGrall-Hill Book Company.
- Wikipedia. (2012). *Mathcad*. Retrieved 2012, from <http://en.wikipedia.org/wiki:>  
<http://en.wikipedia.org/wiki/Mathcad>
- Xavier Jr., A. (2012). *Descritivo da metodologia a ser utilizada e template de relatório final*. Campinas, SP.: Relatório Interno Centro von Braun.