



INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO
DO CONSUMO DE COMBUSTÍVEL DE UMA
EMPRESA RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS**

A carreira, o motorista e a viatura

Rodrigo Manuel da Silva Vargas

**Dissertação para obtenção do Grau de Mestre em
Engenharia Civil**

Júri

Presidente: Professor José Álvaro Pereira Antunes Ferreira
Orientador: Professor João António de Abreu e Silva
Vogal: Professor Filipe Manuel Mercier Vilaça Moura

Outubro 2011

Agradecimentos

Ao Prof. Dr. João António de Abreu e Silva pela orientação que me disponibilizou ao longo do trabalho e pelo tema proposto para a dissertação, que sempre me despertou enorme interesse.

A Jorge Crispim e João Cunha, entre outros, da *Rodoviária de Lisboa S.A.*, pelo constante apoio e disponibilidade quer para fornecerem toda a informação de que necessitei para a realização desta tese quer para o esclarecimento das dúvidas que tive, sempre com grande simpatia e manifesto interesse nos meus desenvolvimentos.

Ao Prof. Dr. Alexandre Bacelar Gonçalves na introdução aos sistemas de informação geográfica e ao software ArcGis, uma vez que, mesmo não o tendo chegado a utilizar os seus conselhos, a sua disponibilidade foi sempre total.

Aos familiares e amigos que sempre me apoiaram e sempre se mostraram interesse quer no motivo da minha dissertação quer no seu desenvolvimento, contribuindo, dessa forma, para a minha motivação.

Ao João Santos, meu primo, pelo seu inestimável apoio à minha introdução ao programa de gestão de base de dados *MySQL*, no qual não tinha quaisquer bases. O seu apoio foi fundamental para o meu sucesso no trabalho de gestão de todos os dados recolhidos, tendo sido incansável no ajuda que me prestou sempre que solicitado.

Aos meus pais pela educação, valores e suporte que sempre foram em toda a minha vida e formação, um sincero agradecimento.

Finalmente à Filipa Leite Rosa, melhor amiga e namorada, pela motivação que sempre me deu, pelo incansável auxílio em múltiplas tarefas e por todo o tempo que isso lhe custou. O seu apoio foi fundamental, nomeadamente nas fases de maior cansaço da minha parte. Por tudo isto e por eu não ter conseguido ser a ajuda que ela foi para mim, um muito obrigado.

Resumo

A definição dos parâmetros que estão envolvidos no consumo de combustível no serviço prestado por uma empresa de transporte de passageiros consiste numa possível ferramenta, complexa mas de grande utilidade, para se otimizar os recursos da própria empresa. O desenvolvimento de uma metodologia que possa fornecer este tipo de informação é um passo importante para a criação de serviços de transporte mais sustentáveis. Para se desenvolver tal metodologia resolveu-se utilizar modelos de regressão múltipla stepwise, para cada um dos principais actores de um serviço de transporte de uma transportadora de passageiros rodoviária: carreiras, motoristas e viaturas.

Com o intuito de realizar estes modelos, foram recolhidos dados relativos à actividade da empresa nos meses de Outubro e Novembro 2009 e março Maio e Setembro de 2010, sendo os 2 últimos utilizados apenas para validação de resultados. A dimensão das amostras recolhidas para o desenvolvimento de cada modelo consistiu em 87 carreiras, 488 motoristas e 105 viaturas. A base de dados recolhida permitiu descrever, com precisão, cada um destes vectores de eficiência energética, tendo sido, então, adaptada a cada um dos modelos.

De um modo geral, as principais variáveis seleccionadas para explicar o consumo de combustível foram o tipo de viatura, a velocidade comercial, a percentagem de percurso com inclinação superior a 5% e a localização das paragens de autocarro. Além destas, foram ainda seleccionadas variáveis relativas a eventos de condução, como o carregamento do ar em excesso de rotação, as desacelerações longitudinais bruscas e as rotações acima dos limites estipulados.

Palavras-chave: Consumo, combustível, carreira, motorista, viatura, eficiência energética, velocidade comercial, regressão linear múltipla

Abstract

The definition of parameters that take part on the fuel consumption of the service provided by transportation companies is a complex but useful tool to optimally manage the companies' resources. The development of a methodology that can provide this information is an important step for developing more sustainable services of transport.

To develop such a methodology were used stepwise multiple regression models for each of the main actors in the service of a company of road transportation: bus lines, drivers and vehicles. For this purpose was gathered data confined to the months of October 2009, November 2009, March 2010, May 2010 and September 2010, with the last two being for validation purposes only. The dimension of the data samples collected for developing the three models consisted in 87 routes, 488 drivers and 105 vehicles. The collected data permitted to describe, with precision, each of the main actors, which were then adapted to each model.

As conclusion, the main variables detected as influential to fuel consumption are the type of vehicle, the commercial speed, the slope greater than 5% and the localization of bus stops. Besides these variables, there were some driving events also selected for being harmful to the fuel's consumption such as the air loading with excessive rotation, the sudden longitudinal decelerations and the rotations above the stipulated maximum value.

Keywords: consumption, fuel, route, driver, bus, energy efficiency, commercial speed, multiple linear regression.

Simbologia e Abreviaturas

Alfabeto romano

Símbolo	Definição
AvrgLenght	- comprimento médio de percurso
AvrLenght	- comprimento médio de percurso
b'_k	- coeficiente de regressão standardizado
ComSpeed	- velocidade comercial média
ComSpeed	- velocidade comercial
D	- estatística de teste do teste de Kolmogorov-Smirnov
DfFits	- número de erros-padrão em que o valor de y_j será alterado se a observação multivariada j for removida da análise.
D_{tabela}	- valor crítico da estatística de teste do teste de Kolmogorov-Smirnov
e	- elasticidade
Ev0140	- percentagem de ocorrência do evento 1040 (carregamento do ar em excesso de rotação)
Ev1007	- percentagem de ocorrência do evento 1007 (desacelerações longitudinais bruscas)
Ev1067	- percentagem de ocorrência do evento 1067 (rotações acima do valor máximo estipulado)
F(X)	- distribuição da variável X
$F_0(X)$	- distribuição específica de X0
H0	- hipótese nula dos testes de hipóteses
H1	- hipótese não nula dos testes de hipóteses
ISup5%	- percentagem percursos com inclinação superior a 5%
$(ISup5\%)^2$	- quadrado da percentagem de percursos com inclinação superior a 5%
k	- número de variáveis independentes (Tabela 4 - ANOVA)
$\log(\text{ComSpeed})$	- logaritmo da velocidade comercial
$\log(\text{DriversAge})$	- logaritmo da idade média de motoristas a circular
$\log(\text{MaxDistStops})$	- logaritmo da distância máxima entre paragens
$\log(\text{VeicMass})$	- logaritmo da massa
N	- dimensão da amostra
N	- distribuição normal
NMonit	- número médio de monitorizações de motoristas

p-value ou p	- nível descritivo
QME	- quadrado médio dos erros;
QMF	- quadrado médio do factor;
R	- coeficiente de correlação múltiplo
R ²	- coeficiente de determinação
R ² _{ajustado}	- coeficiente de determinação ajustado
v	- número de variáveis independentes
S _{xk}	- desvio padrão da variável independente k
S _y	- desvio padrão da variável dependente
SQE	- soma dos quadrados dos erros ou resíduos
SQF	- soma dos quadrados do factor
SQT	- soma dos quadrados totais
variável sdr_1	- variável relativa aos resíduos <i>studentized deleted</i> de cada modelo
variável res_1	- variável relativa aos resíduos obtidos de cada modelo
VeicArtic	- percentagem de percursos feitas com veículo de tipologia Articulado
VeicMidi	- percentagem de condução de viaturas de tipologia Midi
VeicMini	- percentagem de percursos feitas com veículo de tipologia Midi
VeicStand	- percentagem de condução de viaturas de tipologia Standard
VIF	- factor de inflação da variância
X	- variável aleatória
X	- variável independente
x _k	- variável independente
y	- variável dependente
Y	- variável dependente

Alfabeto grego

Símbolo	Definição
α	- nível de significância
β_k	- coeficiente de regressão parcial da equação (3.1)
ε	- resíduo
μ	- média da distribuição normal
σ	- desvio-padrão da distribuição normal

Abreviaturas

IMTT	- Instituto da Mobilidade e dos Transportes Terrestres
PIB	- Produto Interno Bruto
UE	- União Europeia
ITS	- <i>Intelligent Transportation Systems</i>
RL	- Rodoviária de Lisboa S.A.
PNAC	- Plano Nacional para as Alterações Climáticas
PNAEE	- Plano Nacional de Acção para a Eficiência Energética
ENDS	- Estratégia Nacional de Desenvolvimento Sustentável
CAT	- Centro de Actividade de Transportes
GISFRot	- Gestão Integrada de Sistemas de Frotas de Passageiros
FM200	- <i>Fleet manager</i>
SPSS	- <i>Statistical Package for the Social Sciences</i>
ANOVA	- <i>Analysis of Variance</i>

Índice

1.	Introdução.....	1
1.1	Organização actual do sistema de transportes.....	1
1.2	Peso do sector dos transportes no consumo de combustíveis e nas emissões.....	2
1.3	Relação entre os vários modos e os consumos de combustível.....	4
1.4	Necessidades de intervenção no sector dos transportes.....	7
1.4.1	Suportes para um sistema de transportes sustentável.....	7
1.4.4	O IMTT e o futuro dos transportes em Portugal.....	8
1.4.5	O PNAC e o PNAEE.....	8
1.5	A Rodoviária de Lisboa S.A.....	9
1.6	Definição da problemática.....	11
1.7	Objectivos e âmbito da investigação.....	12
2.	Estado da arte.....	13
2.1	O tipo de veículo.....	13
2.1.1	Massa dos veículos.....	13
2.2	A caracterização de carreira.....	13
2.2.1	Caracterização de paragens de autocarro.....	13
2.2.2	Velocidade comercial.....	14
2.3	A infra-estrutura e os condicionalismos de funcionamento.....	15
2.3.1	Caracterização topográfica.....	15
2.4	O motorista.....	15
2.4.1	Idade média de motoristas.....	15
2.5	A política da empresa de transporte.....	16
2.5.1	Eventos de condução.....	16
2.5.2	Número médio de monitorizações formativas de motoristas.....	18
3.	Conceitos estatísticos sobre Regressão Linear.....	21
3.1	A regressão linear múltipla.....	21

3.2	Coeficientes de regressão não standardizados.....	22
3.3	Coeficientes de regressão standardizados.....	22
3.4	Métodos de regressão linear múltipla	23
3.5	Pressupostos de um modelo de regressão linear	23
3.5.1	Linearidade.....	24
3.5.2	Variância constante dos resíduos - Homocedasticidade.....	24
3.5.3	Independência dos resíduos	24
3.5.4	Normalidade da distribuição dos resíduos	24
3.6	Multicolinearidade e singularidade	25
3.7	Tolerância e factor de inflação da variância (VIF)	25
3.8	Dimensão da amostra	26
3.9	Outliers	26
3.10	Variáveis <i>Dummie</i>	26
3.11	Transformação de variáveis.....	27
3.12	Metodologias de análise de um modelo de regressão linear	28
3.13	Verificação da adequabilidade da regressão ajustada.....	28
3.14	Análise de variância do modelo de regressão linear	29
3.15	Verificação dos pressupostos de linearidade e normalidade	29
3.16	Testes paramétricos.....	31
3.16.1	Teste Kolmogorov-Smirnov	31
3.16.2	Teste Shapiro-Wilk.....	31
3.17	Estimação de elasticidades	32
4.	Descrição das amostras estatísticas e desenvolvimento das variáveis	33
4.1	Descrição da amostra de Carreiras	33
4.1.1	Consumo	34
4.1.2	Tipologia de carreira	34
4.1.3	Inclinação	35
4.2	Descrição da amostra de Motoristas	36

4.2.1 Consumo	36
4.2.2 Idade de motoristas	37
4.2.3 Número de monitorizações formativas	37
4.3 Descrição da amostra de Viaturas	38
4.3.1 Consumo	38
4.3.2 Tipologia de viaturas.....	39
4.3.3 Idade de viaturas.....	39
4.4 Descrição dos processos de criação de variáveis	40
4.4.1. Tipo de veículo	41
4.4.2. Caracterização da carreira.....	43
4.4.3. A infra-estrutura e condicionalismos de funcionamento	45
4.4.4. O motorista.....	46
4.4.5. A política da empresa de transporte	48
4.4.5.1 Número médio de monitorizações formativas de motoristas.....	48
4.5 Limitações na recolha de dados	49
4.5.1 Consumo de combustível médio	49
4.5.2 Variáveis de inclinação	50
4.5.3 Número de passageiros transportados	50
4.5.4 Amostra no modelo de veículos	50
5. Apresentação e descrição de resultados.....	53
5.1 Resultados do modelo de carreiras	53
5.1.1 Sumário do modelo obtido	53
5.1.2 Diagnósticos de multicolinearidade	55
5.1.3 Análise de resíduos	56
5.1.4 Testes de normalidade	56
5.2 Resultados do modelo de motoristas.....	57
5.2.1 Sumário do modelo obtido	57
5.2.2 Diagnósticos de multicolinearidade	59

5.2.3	Análise de resíduos	59
5.2.4	Testes de normalidade	59
5.3	Resultados do modelo de viaturas	60
5.3.1	Sumário do modelo obtido	60
5.3.2	Diagnósticos de multicolinearidade	62
5.3.3	Análise de resíduos	62
5.3.4	Testes de normalidade	62
6.	Discussão e interpretação de resultados	65
6.1	Modelo Carreiras	65
6.1.1	Considerações gerais	65
6.1.2	Análise de variáveis	66
6.2	Modelo Motoristas	72
6.2.1	Considerações gerais	72
6.2.1	Análise de variáveis	73
6.2.2	Validação do modelo de Motoristas	77
6.3	Modelo Viaturas	78
6.3.1	Considerações gerais	78
6.3.2	Análise de variáveis	78
6.3.3	Validação do modelo de Viaturas	82
7.	Conclusões	85
7.1	Principais resultados	86
7.1.1	Tipologia de viaturas	86
7.1.2	Velocidade comercial	86
7.1.3	Monitorizações formativas de motoristas	87
7.1.4	Idade motoristas	87
7.1.5	Eventos de condução	87
7.2	Outras medidas de conservação de energia	88
7.3	Limitações de metodologia desenvolvida	88

7.4 Trabalhos futuros	89
Referências bibliográficas.....	91

Índice de Tabelas

Tabela 1 – Evolução dos eventos de Conforto.....	16
Tabela 2 – Evolução dos eventos de Segurança	16
Tabela 3 – Evolução dos eventos de Desempenho ambiental.....	17
Tabela 4 – Tabela de análise de variância (ANOVA)	29
Tabela 5 – Sumário de elasticidades face ao tipo de transformação das variáveis	32
Tabela 6 – Sumário do modelo de Carreiras.....	53
Tabela 7 – Coeficientes estatísticos das variáveis do modelo de Carreiras	55
Tabela 8 – Sumário do modelo de Motoristas.....	57
Tabela 9 – Coeficientes estatísticos das variáveis do modelo de Motoristas	58
Tabela 10 – Sumário do modelo de Viaturas.....	60
Tabela 11 – Coeficientes estatísticos das variáveis do modelo de Viaturas	61
Tabela 12 – Correlação entre Consumo de combustível e quadrado da inclinação superior a 5% com os Centros de Actividade de Transporte	69
Tabela 13 - Tabela de inclinação superior a 5% e consumo de combustível por Centro de Actividade de Transporte.....	70

Índice de Figuras

Figura 1- Evolução do número de transporte de passageiros entre 1995 e 2006.....	1
Figura 2 – Relação entre o número de automóveis por 1000 habitantes e o PIB per capita expresso em dólares norte-americanos de 1988. Dados de 17 países da Europa e dos EUA.....	3
Figura 3 – Emissões de CO ₂ , na União Europeia, por modo de transporte.....	3
Figura 4 – Evolução das emissões de GEE no sector dos transportes (1990-2020).....	4
Figura 5 – Custos externos médios por modo de transporte ao nível do transporte de passageiros, em 1995, na União Europeia.....	5
Figura 6 – Bilhões de passageiros-quilómetro por modo de transporte, na União Europeia.	6
Figura 7 – Consumo (L/100km) vs Velocidade comercial	14
Figura 8 – Variação do factor de emissão de CO ₂ face à variação da velocidade de circulação.	17
Figura 9 – Consumo de combustível vs Aceleração vs Velocidade	18
Figura 10 – Evolução do consumo específico de gasóleo.....	19
Figura 11 – Consumo de combustível médio para a amostra de carreiras recolhida.	34
Figura 12 – Número de carreiras por tipologia para a amostra recolhida.....	35
Figura 13 – Percentagem de inclinação para diferentes patamares de inclinação da amostra recolhida.	35
Figura 14 – Consumo de combustível médio para a amostra de motoristas recolhida.	36
Figura 15 – Idade de motoristas para a amostra recolhida.	37
Figura 16 – Número de monitorizações por motorista para a amostra recolhida.	37
Figura 17 – Consumo de combustível de viaturas para a amostra recolhida.	38
Figura 18 – Número de viaturas por tipologia para a amostra recolhida.	39
Figura 19 – Idade de viaturas para a amostra recolhida.....	40
Figura 20 – Velocidade comercial vs Distância média entre paragens.....	67
Figura 21 – Percentagem de ocorrência do evento 1067 vs Número médio de monitorizações formativas de motoristas	68
Figura 22 – Velocidade comercial (km/h) vs Logaritmo da distância máxima entre paragens.....	71
Figura 23 – Comparação entre valores previstos de logaritmo de consumo pelo modelo de Carreiras e valores observados da amostra de validação.	72

Figura 24 – Percentagem de percursos com inclinação superior a 5% vs Velocidade comercial.	75
Figura 25 – Percentagem de ocorrência do evento 1060 vs Comprimento médio de percurso.	76
Figura 26 - Comparação entre valores previstos de logaritmo de consumo pelo modelo de Motoristas e valores observados da amostra de validação.....	77
Figura 27 – Velocidade comercial vs Comprimento médio de percurso.....	80
Figura 28 – Percentagem de ocorrência do evento 1007 vs Logaritmo da idade média de motoristas.	81
Figura 29 – Percentagem de ocorrência do evento 1007 vs Logaritmo da idade média de motoristas.	82
Figura 30 - Comparação entre valores previstos de logaritmo de consumo pelo modelo de Viaturas e valores observados na amostra de validação.	83

1. Introdução

1.1 Organização actual do sistema de transportes

A sociedade está cada vez mais dependente de transportes seguros, confortáveis, fiáveis e rápidos, quer para ir trabalhar, como para viajar e para distribuir produtos, que determinam o nosso estilo de vida. Assim, é essencial que o sistema de transportes seja eficiente para manter a competitividade da economia. Os transportes constituem, actualmente, um sector económico que representa cerca de 10% da riqueza europeia em termos de produto interno bruto. Esta indústria vale cerca de um trilião de euros por ano e emprega mais de 10 milhões de pessoas (CE, 2003).

Nas últimas décadas tem-se observado uma grande evolução no sector dos transportes, uma vez que os desenvolvimentos na política de transportes dos últimos 50 anos foram bem sucedidos, tendo conseguido ganhos na facilidade, preço e segurança da movimentação de pessoas e bens. Uma grande fatia desse sucesso do desempenho global do sistema de transportes pode ser associada ao modo rodoviário, tendo-se tornado dominante nos últimos anos no suporte aos movimentos de pessoas e bens, o que se pode dever à independência e flexibilidade que proporciona (VIEGAS, 2003).

As tendências no desempenho dos transportes, nomeadamente no transporte de mercadorias, acompanham a evolução económica. Enquanto o produto interno bruto (PIB, medido a preços constantes de 1995) cresceu a uma taxa média anual de 2,4% de 1995 a 2006, o desempenho do transporte de mercadorias, medido em toneladas-quilómetro, cresceu em 2,8% ao ano. Ainda durante esse período, o desempenho do transporte de passageiros, medido em passageiros-quilómetro, cresceu a uma taxa média anual de 1,7% (EUROSTAT, 2009).

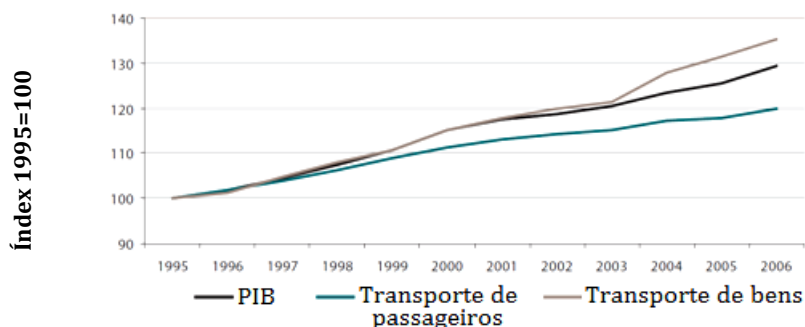


Figura 1- Evolução do número de transporte de passageiros entre 1995 e 2006 (index 1995=100). (Fonte: EUROSTAT, 2009)

Apesar de, até hoje, os transportes nunca terem sido tão rápidos e baratos, ainda são detectáveis situações desadequadas, nomeadamente ao observar-se que o transporte continua a sofrer com congestionamento, poluição, acidentes, preços inapropriados, sub-investimento, e um défice continuado de uma cadeia de transportes genuinamente lógica (SHORT, 2003).

1.2 Peso do sector dos transportes no consumo de combustíveis e nas emissões

Hoje em dia, o petróleo continua a ser um dos principais combustíveis utilizados nos meios de transporte e prevê-se que continue assim durante décadas, dadas as suas qualidades inerentes de elevada densidade energética e um teor de CO₂ muito inferior ao do carvão. No entanto, como já foi referido anteriormente, o petróleo é um recurso finito e muito procurado, relativamente ao qual a União Europeia depende, essencialmente, da importação. Os transportes representam cerca de 70% do consumo anual de petróleo na União Europeia e cerca de 96% dos veículos a motor têm, por base, o petróleo (GV, 2006). É fundamental que se tomem as medidas necessárias para diminuir os consumos registados actualmente ao nível dos combustíveis fósseis. Mesmo o progresso que tem vindo a ser desenvolvido ao nível da eficiência dos motores dos automóveis não é suficiente para compensar o rápido crescimento do número de quilómetros percorridos anualmente por pessoa (ANDRÉ, 2008).

Em Portugal, o consumo de energia pelo sector dos transportes tem um peso de 39% da totalidade do consumo de energia em todos os sectores, sendo que o transporte rodoviário, em particular, tem um peso de 87% do sector. Além do mais, estima-se que, durante os últimos 15 anos, o número de carros em circulação tenha crescido cerca de 135% (IMTT, 2008).

Pode dizer-se que a principal razão para o contínuo crescimento da procura de transporte, no caso dos passageiros, é o constante desenvolvimento da motorização (CE, 2003). No entanto, é necessário ter em conta que o sector dos transportes não pode continuar a crescer como até aqui. Actualmente, os congestionamentos já afectam os sectores rodoviário, aéreo e ferroviário e resultam em demoras, desconforto, aumento da poluição, bem como em custos mais elevados e menor qualidade dos próprios serviços. O congestionamento dos transportes rodoviários leva ainda a um aumento da poluição do ar nos centros urbanos até níveis ameaçadores para a saúde pública. A tendência de todas as famílias terem, pelo menos, um automóvel particular, pode considerar-se um dado adquirido. A correlação estatística entre o PIB *per capita* e a taxa de motorização é evidente e encontra-se demonstrada na Figura 4:

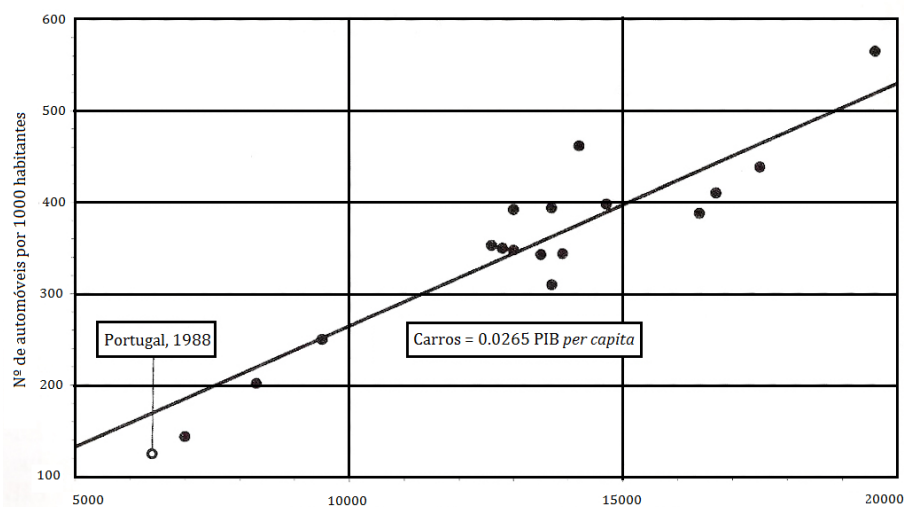


Figura 2 – Relação entre o número de automóveis por 1000 habitantes e o PIB per capita expresso em dólares norte-americanos de 1988. Dados de 17 países da Europa e dos EUA. (Fonte ANDRÉ, 2008)

Em 1990, a taxa de motorização em Portugal rondava os 215 veículos ligeiros por cada 1000 habitantes. Em 2005, esta taxa atingiu cerca de 475 veículos por cada 1000 habitantes (PNAC, 2006).

A situação para a qual se está a caminhar, ao nível do consumo energético, está a tornar-se, portanto, insustentável. Apesar da redução desse consumo ao nível da indústria, e consequentemente da diminuição das suas emissões de CO₂, e dos sectores terciário e doméstico aumentarem apenas ligeiramente o consumo, o sector dos transportes continua a consumir e a poluir cada vez mais. Este elevado consumo energético registado no sector dos transportes tem então, como consequência, um grande peso nas emissões de CO₂, como se pode verificar nas Figuras 2 e 3.

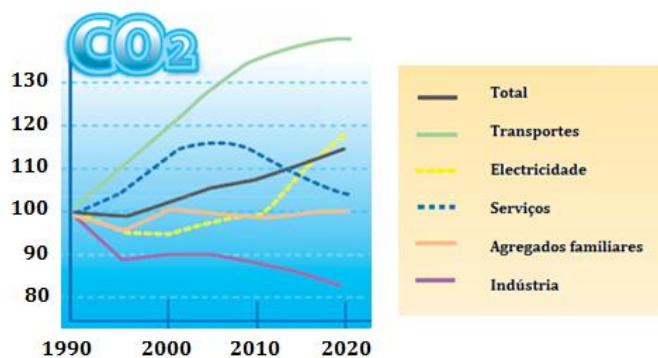


Figura 3 – Emissões de CO₂, na União Europeia, por modo de transporte (Índice de base - 100 em 1990). (Fonte CE, 2003)

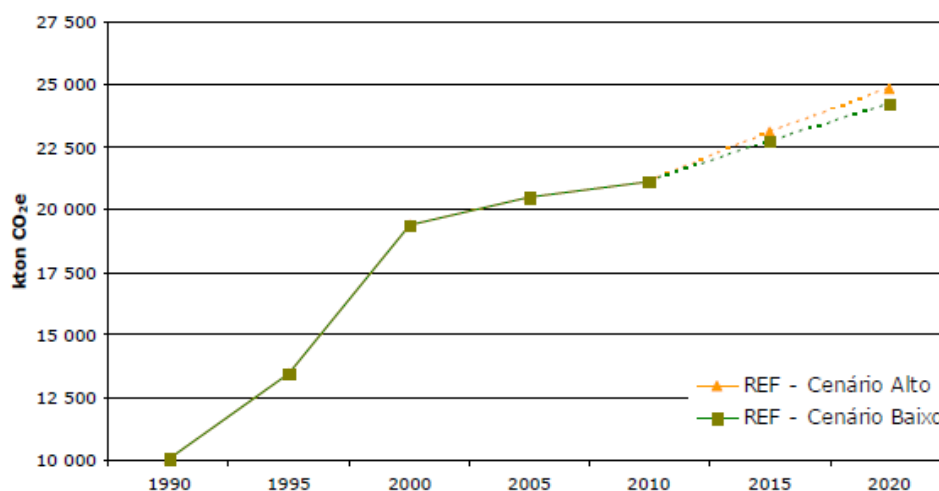


Figura 4 - Evolução das emissões de GEE no sector dos transportes (1990-2020). (Fonte PNAC, 2006)

Assim, é extremamente importante que se concentrem esforços com o intuito de se criarem as medidas necessárias para se diminuir as emissões gasosas, como estabelecido no protocolo de Quioto.

1.3 Relação entre os vários modos e os consumos de combustível

A predominância da estrada é muito acentuada no caso do transporte de passageiros, como se pode constatar pela Figura 4, em que o rodoviário representa cerca de 74% da quota de mercado, enquanto que a via aérea, com 8%, já ultrapassou o caminho-de-ferro, que não sai dos 6%. A grande força do transporte rodoviário face aos restantes consiste na competitividade resultante da sua flexibilidade inigualável e no seu baixo preço (CE, 2001).

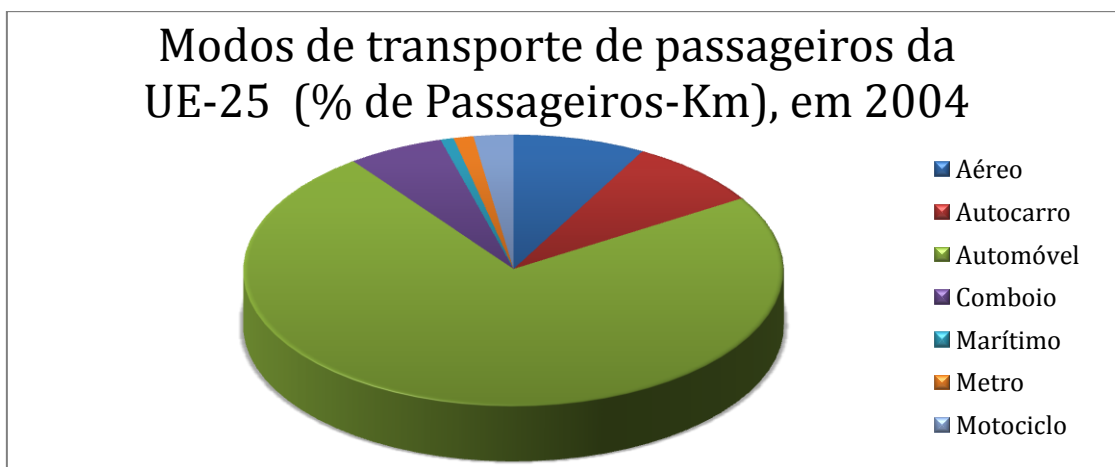


Figura 4 - Modos de transporte de passageiros da UE-25, em 2004. (Fonte CE, 2006b)

Pela figura anterior, constatou-se que existe uma clara predominância na procura pelo transporte aéreo mas principalmente pelo rodoviário. No entanto, é a estes modos de transporte que estão associados os maiores custos externos.

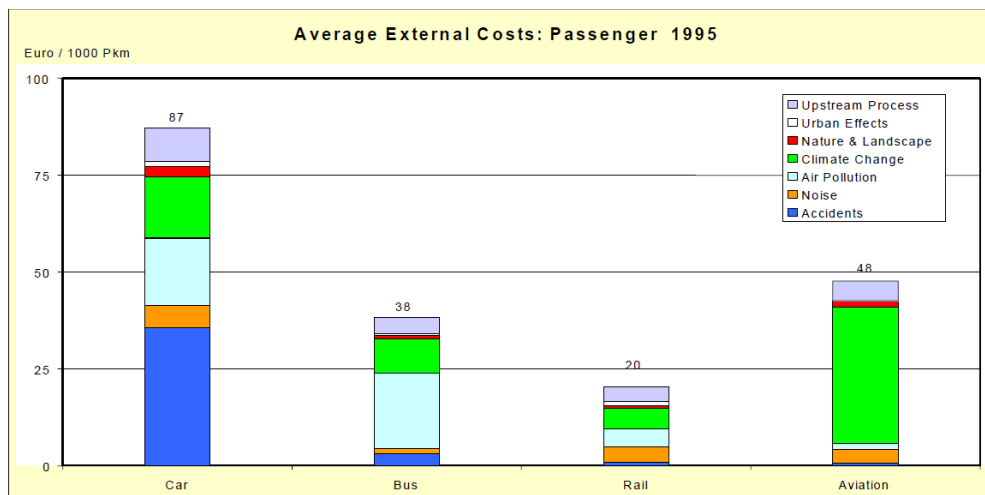


Figura 5 – Custos externos médios por modo de transporte ao nível do transporte de passageiros, em 1995, na União Europeia. (Fonte EEA, 2001)

É natural que, à medida que a vida das pessoas se enriquece com novas actividades e interesses, as deslocações tendam a aumentar, mas esse progresso acaba por ter algumas consequências negativas se não for planeado à escala colectiva. Os automóveis tornam as cidades menos habitáveis se não forem oferecidos outros meios de transporte competitivos (ANDRÉ, 2008).

É então necessário enfrentar o crescente desequilíbrio que actualmente existe na procura entre os diferentes modos de transporte com o acentuado crescimento do transporte rodoviário individual e transporte aéreo de passageiros e a ainda a estabilização da procura dos restantes meios de transporte, como se pode verificar na Figura 6. Para fazer frente a este problema, no Livro Branco dos transportes (CE, 2011) definiu-se como meta o assegurar de uma concorrência regulada entre modos de transporte, associando os diferentes modos, a fim de se alcançar a intermodalidade.

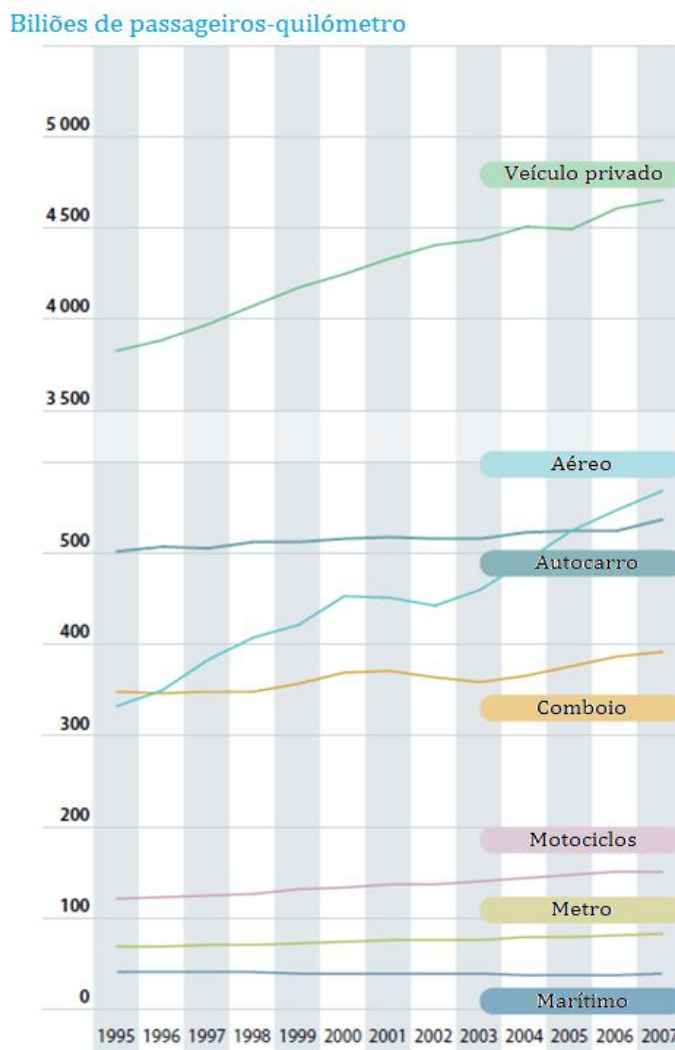


Figura 6 – Biliões de passageiros-quilómetro por modo de transporte, na União Europeia. (Fonte CE, 2002)

Nos dias de hoje, o sector dos transportes vive uma situação de congestão crónica, em que 7 500 km de estradas, que representam cerca de 10% da rede, estão diariamente afectados por congestionamentos, cerca de 16 000 km de vias férreas, 20% da rede, podem ser considerados zonas de estrangulamento e 16 dos principais aeroportos da União Europeia registaram atrasos superiores a um quarto de hora em mais de 30% dos seus voos (CE, 2001). O congestionamento implica grandes riscos ao nível da perda de competitividade da economia europeia, uma vez que aumentam o consumo de combustível e, conseqüentemente, a poluição. Grande parte das emissões gasosas provêm dos veículos rodoviários e embora, na generalidade dos sectores, as emissões de CO₂ estejam a diminuir, as emissões relativas aos transportes ainda estão a aumentar (CE, 2003). A poluição atmosférica e sonora intensifica-se todos os anos e a circulação urbana está na origem de 40% das emissões de CO₂ e de 70% das emissões de outros poluentes resultantes dos transportes rodoviários (CE, 2007).

1.4 Necessidades de intervenção no sector dos transportes

1.4.1 Suportes para um sistema de transportes sustentável

Actualmente, é necessário redefinir o transporte público de outra forma, e esse caminho tem de consistir em suplantar o automóvel, oferecer melhores condições de transporte e a um preço inferior. O objectivo da política europeia de transportes é criar um sistema de transportes sustentável que satisfaça as necessidades económicas, sociais e ambientais da sociedade e conduza a uma sociedade sem exclusões e a uma Europa plenamente integrada e competitiva. A tendência aponta para a necessidade de satisfazer um aumento da procura de acessibilidade num contexto de preocupações crescentes ao nível da sustentabilidade. As prioridades mais imediatas parecem ser a melhoria da integração dos diversos modos de transporte, como forma de melhorar a eficiência global do sistema, e ainda a aceleração do desenvolvimento e da implantação de tecnologias inovadoras, baseando-se numa lógica que coloca os utentes e os trabalhadores do sector dos transportes, as suas necessidades e os seus direitos, no centro do processo político (CE, 2009).

Segundo T&E, 2004, um transporte sustentável "deverá implicar uma utilização dos meios de transporte que satisfaça as necessidades presentes sem prejudicar as gerações vindouras". Para isto, é necessário que se limite o acesso e a utilização do automóvel, pela promoção da circulação pedonal ou através de medidas políticas de ordenamento de território e transportes adequadas. Naturalmente, é fundamental que se criem as infra-estruturas necessárias mas também é fundamental que se desenvolvam campanhas de sensibilização e de informação a estas novas medidas (BORREGO, 2005).

A crescente utilização das tecnologias de informação com sistemas de apoio à exploração e de posicionamento global para além da gestão integrada dos sistemas de transportes, da "bilhética sem contacto" (actualmente já implementado em algumas empresas de transporte em Lisboa), do tacógrafo digital e das caixas negras, são aspectos que hoje já são tidos em conta no sistema de transportes colectivos (PROPLANO, 2006).

Pela constatação de uma degradação generalizada da qualidade de vida dos cidadãos europeus, afectados pelo crescente congestionamento das cidades, a Comissão propõe-se a incentivar a permuta das boas práticas, com vista a uma melhor utilização dos transportes colectivos e das infra-estruturas em que estes actuam (CE, 2001). É então necessária uma maior atenção por parte dos poderes públicos locais para conciliar a modernização do serviço público com a racionalização do uso do veículo particular. Medidas como estas são fundamentais para se alcançar um desenvolvimento sustentável mas, logicamente, são difíceis de aplicar. Ainda assim, para que se

possa respeitar os compromissos internacionais assumidos em Quioto, no sentido da redução das emissões poluentes, este é o caminho que tem de ser seguido.

1.4.4 O IMTT e o futuro dos transportes em Portugal

Em Portugal, foi criado, em 2007, o Instituto da *Mobilidade e dos Transportes Terrestres (IMTT)* que tem como missão regular, fiscalizar e exercer funções de coordenação e planeamento do sector dos transportes terrestres, bem como supervisionar e regular as actividades deste sector. Como tal, o *IMTT* propôs-se a desenvolver medidas com o intuito de minorar as emissões de carbono e de promover a intermodalidade, otimizando o desempenho global dos modos de transporte público, visando incrementar a sua utilização e reduzir o congestionamento gerado pelo transporte individual. Com essas metas definidas, o *IMTT* tem vindo a promover a introdução de veículos mais amigos do ambiente nas frotas de transporte público, como, por exemplo, a introdução com sucesso de veículos movidos a gás natural em Lisboa e Porto e ainda uma experiência piloto com mini autocarros eléctricos em várias cidades bem como o desenvolvimento do *Plano Nacional de Plataformas Logísticas*. Estas experiências são fundamentais para cumprir os objectivos de alcançar a intermodalidade, de modo a alterar a repartição modal, reduzindo assim a quota do transporte rodoviário (IMTT, 2008).

Neste tipo de projectos insere-se o *GisFrot*, fruto da colaboração entre o IMTT e a *Rodoviária de Lisboa S.A.*, e que consiste num programa de gestão de frotas, promovendo acções de eco-condução e de monitorização formativa dos seus condutores.

1.4.5 O PNAC e o PNAEE

No período entre 1990 a 2005, o sector dos transportes caracterizou-se pelo aumento do consumo energético em cerca de 102%, correspondente a uma taxa anual de crescimento de 4,8%. O principal agente responsável por este aumento foi o modo rodoviário, que registou um crescimento do consumo de energia de 107%. Neste modo de transporte destacam-se as deslocações efectuadas em transporte individual, que cresceram mais de 111%, a um ritmo médio de 5,1% por ano (PNAC, 2006).

O grande contributo dos transportes ao nível das emissões de gases causadores de efeito de estufa obrigou a que se desenvolvessem relatórios que pretendem criar medidas concretas com o intuito de alcançar as metas traçadas pelo protocolo de Quioto. Desta forma, foram desenvolvidos o PNAC (Plano Nacional para as Alterações Climáticas) e o PNAEE (Plano Nacional de Acção para a

Eficiência Energética).

Actualmente, definiu-se como meta global da UE que se atinjam padrões de produção e consumo sustentáveis. Pretende-se, então, separar a relação entre o aumento do crescimento económico, em cada um dos sectores de actividade, dos impactos negativos no ambiente resultantes da utilização dos recursos naturais. O Programa Nacional para as Alterações Climáticas (PNAC) e a Estratégia Nacional de Desenvolvimento Sustentável (ENDS) assim como as novas metas de combate às alterações climáticas têm o intuito de criar as condições para o cumprimento destes objectivos (REA, 2007).

Para que Portugal consiga atingir as metas estabelecidas no protocolo de Quioto terá que seguir medidas de redução da emissão dos GEE (gases do efeito de estufa). O PNAC tem, então, como objectivo, controlar e reduzir as emissões de gases com efeito estufa, respeitando os compromissos estabelecidos para Portugal. Foram então tomadas medidas tidas como adequadas para que o país conseguisse atingir os objectivos definidos. Já o PNAEE consiste num plano composto por um conjunto de programas e medidas de eficiência energética, num horizonte temporal que se estende até ao ano de 2015. No sector dos transportes, o PNAEE apresenta três programas que são: *Renove Carro*, *Mobilidade Urbana* e ainda *Sistema de Eficiência Energética Transportes*. Algumas das medidas presentes no PNAC, relativas ao sector do transporte colectivo de passageiros, apresentam-se, em seguida:

- Incentivos à transferência modal através do aumento da qualidade e diversidade da oferta no transporte colectivo de passageiros;
- Programa de incentivo ao abate de veículos em fim de vida;
- Ampliação das frotas de veículos a gás natural na CARRIS e STCP;
- Incentivo a acções de formação de sensibilização a condutores profissionais para uma condução eco-eficiente.

1.5 A Rodoviária de Lisboa S.A.

A Rodoviária de Lisboa, S.A. foi criada em 1991, no quadro da cisão da Rodoviária Nacional E.P, a partir da nacionalização dum universo de 92 empresas estabelecidas como Operadores de Transporte Colectivo Rodoviário de Passageiros.

Em 1995 ocorreu a privatização e cisão da Rodoviária de Lisboa, S.A, passando a ser detida pelo grupo económico Barraqueiro Transportes tendo operando na Área Metropolitana de Lisboa, nos

concelhos de Loures, Odivelas e Vila Franca de Xira, e está dividida em 3 Centros de Actividades de Transportes (CAT): Caneças (CAT 5), Bucelas (CAT 6) e Santa Iria de Azóia (CAT 7). Com um universo de 375 viaturas e de cerca de 773 trabalhadores, a Rodoviária de Lisboa opera ao longo de 1300 Kms de rede concessionada, distribuídos por 94 carreiras. As viaturas são de diferentes tipologias, marcas e modelos, e têm uma idade média de 14,8 anos. Todas as viaturas da frota operacional da RL utilizam o gasóleo como fonte energética. A sua oferta garante a realização de carreiras Directas, Urbanas, Suburbanas, Praias (serviço sazonal) e ainda alugueres fixos e ocasionais. Desta forma, a Rodoviária de Lisboa S.A. contribui, de forma marcante, para a concretização do desígnio da mobilidade e da qualidade de vida das populações da zona oriental da Área Metropolitana da Região de Lisboa (RL, 2008a).

A RL é uma empresa organizada, no que se refere ao acompanhamento, verificação e à gestão dos seus consumos energéticos. Outro exemplo disso é o projecto GisFrot que vem sendo implementado desde 2004, monitorizando o comportamento da condução dos motoristas. O projecto GISFRot pretende desenvolver a Gestão Integrada de Sistemas de Frotas de Passageiros, visando a optimização energética e a melhoria da qualificação dos motoristas. O projecto foi desenvolvido com o apoio financeiro do IMTT, tendo ainda com parceiros o idMEC (IST), a RODinform e a Multifrota fleet. O sector de transporte rodoviário de passageiros apresenta grandes desafios, sendo imperativa a necessidade de garantir uma mobilidade sustentável, a partir da melhoria da qualificação de motoristas e da qualidade de serviço. Para isso, é necessário que a qualificação dos motoristas contribua para a melhoria da qualidade do serviço e da sua eficiência energética e ambiental.

O GisFrot começou a ser desenvolvido em 2004 e encontra-se actualmente na sua terceira fase. Este projecto obriga um conhecimento detalhado das viaturas (tipologias, caixas e motores) bem como uma boa interacção com o sector de Manutenção. Desta forma é possível obter dados objectivos sobre o desempenho dos motoristas em contexto de trabalho. À posteriori, é ainda realizada uma monitorização formativa dos motoristas, através da análise dos resultados que obtiveram. A informação das viaturas é descarregada diariamente, via wireless, para um computador com a aplicação Fleet Manager (FM), sempre que as viaturas entram nos parques. A comparação do desempenho dos motoristas, independentemente do tempo de condução e da viatura que utilizem, é medida através da taxa de ocorrência do evento respectivo e consiste no quociente entre a duração do evento e o tempo de condução. Os eventos consistem então em condições pré-definidas que se verificam durante a utilização de uma viatura e que são registados no FM.

A grande meta deste programa consiste então em desenvolver um "triângulo" de eficiência energética (composto pelos vectores "Carreira", "Motorista" e "Viatura") que possa, no futuro,

atribuir uma viatura a uma determinada carreira e a um determinado motorista de forma a otimizar consumos, tempos de viagem e conforto dos passageiros.

Uma vez que as viaturas se encontram equipadas com sistemas de GPS é possível, na análise posterior dos eventos, identificar o local preciso em que se registaram. Além disso, é emitido um sinal acústico que permite dar a conhecer ao motorista que foi registado um evento nesse instante. Os dados que são validados são comparados pela hora de ocorrência e pela viatura com o programa de escalas, que indica qual o motorista que nesse momento conduzia a viatura e qual o serviço que efectuava. A análise do desempenho do motorista apresenta os níveis de desempenho, global e por evento, bem como o número de horas de condução e as várias viaturas em que operou. Nesta fase, está estabelecido que todos os motoristas da empresa devem ter pelo menos uma sessão de monitorização formativa por ano. Até agora, o programa GisFrot tem permitido que se alcance o principal objectivo que inicialmente se tinha traçado que era monitorização formativa de todos os motoristas a partir do seu desempenho. Ainda assim, o foram garantidos outros aspectos relevantes como são a obtenção dos consumos associados a determinado tipo de eventos em alguns veículos (apenas os equipados com linha Can) mas também a redução do consumo das viaturas de 2,5% (RL, 2008a).

Actualmente, o vector "Motorista" já está caracterizado com grande precisão, pelo que se espera que as próximas etapas do projecto GisFrot consistam no desenvolvimento dos vectores "Carreira" e "Viatura". É também nesse seguimento que se pretende desenvolver esta dissertação, pelo que se espera que a análise efectuada permita identificar os principais aspectos passíveis de melhoria na cadeia de operações das carreiras e no desempenho do motorista.

1.6 Definição da problemática

Uma das características mais importantes do transporte sustentável é o uso eficiente da energia, ou seja, a capacidade de transportar o máximo de carga gastando o mínimo de combustível. Nesse sentido, as acções preventivas e de orientação de motoristas bem como a maior verificação do desgaste das viaturas são cada vez mais parte do dia-a-dia do transporte rodoviário, nomeadamente em empresas de transporte.

Uma vez que o combustível tem um peso bastante grande na economia de uma empresa de transportes, qualquer empresa deve tomar especial atenção à gestão do mesmo na sua actividade. Actualmente, com o aumento dos custos do barril de petróleo, esta é uma área que, mais do que nunca, importa explorar de forma a otimizar-se a utilização dos recursos de que se dispõe, desde

os veículos aos motoristas. Segundo a empresa KENWORTH (2008), existem 6 áreas principais que influenciam os consumos de combustível por parte das viaturas, que são:

1. Aerodinâmica;
2. Componentes;
3. Tecnologia avançada;
4. Gestão de rotas;
5. Comportamento do condutor;
6. Manutenção adequada.

De forma a aperfeiçoar os conhecimentos sobre quais os intervenientes e de que forma estes influenciam o consumo de combustível de uma viatura, é necessário promover pesquisas científicas com esse intuito.

1.7 Objectivos da investigação

O objectivo desta dissertação consiste na análise dos principais vectores de eficiência energética da operadora *Rodoviária de Lisboa*. Para isso, serão desenvolvidos 3 modelos de regressão linear múltipla de previsão do consumo de combustível de cada um dos vectores de eficiência energética: Carreiras, Motoristas e Viaturas. Através destes modelos serão detectadas as variáveis que melhor conseguem explicar os consumos de combustível dos intervenientes já referidos. Para isso, irá ser feita uma recolha das variáveis relativas à caracterização topográfica das carreiras, à localização das paragens de autocarros, às velocidades comerciais verificadas, aos eventos de condução registados pelo FM200 e à caracterização dos motoristas e das viaturas. A recolha de dados será feita em meses considerados padrão, a nível laboral e escolar, evitando os períodos de férias.

Através da definição destes modelos de previsão, a Rodoviária de Lisboa irá dispor de uma ferramenta que permita fazer uma melhor gestão dos recursos de que dispõe e ainda identificar os principais aspectos passíveis de melhoria na cadeia de operações das carreiras e no desempenho do motorista.

2. Estado da arte

O presente capítulo abrange a fundamentação teórica que serviu de sustentação ao desenvolvimento desta tese. Pretende-se assim fazer um levantamento de estudos acerca de indicadores que se consideram, à partida, como preponderantes na previsão do consumo de combustível de veículos rodoviários de transporte de passageiros.

A execução desta pesquisa permitirá tomar conhecimento dos resultados apresentados por investigações relevantes nesta área. Desta forma, com o presente trabalho, pretende-se verificar alguns dos resultados indicados nesses projectos mas também desenvolver novas conclusões e sugestões para melhorar a sustentabilidade de uma empresa de transporte.

Através da recolha de informação sobre os estudos que foram feitos nesta temática, foi possível detectar alguns factores chave que influenciam o consumo de combustível como o tipo de veículo, a caracterização física de uma carreira ao nível das suas infra-estruturas e condicionamentos de funcionamento, o tipo de condução efectuada pelos motoristas e até a política da empresa, nomeadamente através de acções de formação que para motoristas e sistemas de monitorização da sua actividade.

2.1 O tipo de veículo

2.1.1 Massa dos veículos

O valor da massa das viaturas que cumprem o serviço de transporte de passageiros tem uma grande influência no seu consumo esperado. SIMÕES, (2005), analisou a redução da massa média da frota da operadora *Horários do Funchal*, enquadrada na renovação da frota e, através da redução da massa da frota em 10%, previu-se que a diminuição ao nível do consumo seria cerca de 2%.

2.2 A caracterização de carreira

2.2.1 Caracterização de paragens de autocarro

A caracterização da disposição de paragens de autocarro pode desempenhar um papel relevante quando se pretende avaliar a eficiência energética do serviço de uma transportadora de passageiros. Na verdade, quanto maior for a proximidade entre paragens de uma carreira, mais frequente será o ciclo de “*pára/arranca*” por parte da viatura, o que por sua vez resultará numa

natural diminuição da sua velocidade comercial. Este é um aspecto decisivo na definição do traçado de cada carreira, uma vez que tem bastante influência na procura da carreira.

Relacionando estas variáveis com o consumo de combustível, a VOLVO, 2010, sugere que uma paragem adicional em cada 10 km pode aumentar o consumo de combustível até 35%. Já em SIMÕES, 2005, conclui-se que as paragens efectivas de entrada e saída de passageiros, para a frota da operadora *Horários do Funchal*, representavam 21% dos consumos anuais.

2.2.2 Velocidade comercial

A velocidade comercial consiste num dos principais avaliadores de qualidade de desempenho de qualquer serviço de transporte de passageiros, pelo que terá um papel decisivo na procura, estando então directamente relacionada com os custos para o operador e para o usuário (BARNIOL, 2005). Além disto, a velocidade comercial também poderá ser um indicador da eficiência do consumo de combustível do serviço disponibilizado e é desta forma que se pretende avaliar a sua importância neste trabalho. A velocidade comercial é afectada pela acessibilidade, circulação e estacionamento de viaturas mas também pela distância entre paragens e pela sua frequência (CAMPBELL, 1991). Em UITP, 2009, verificou-se que a velocidade comercial de uma viatura tem um grande efeito sobre o seu consumo, como se pode constatar pela Figura 7. Na verdade, quanto maior é a velocidade comercial média, menor é o consumo de combustível desse deslocamento.

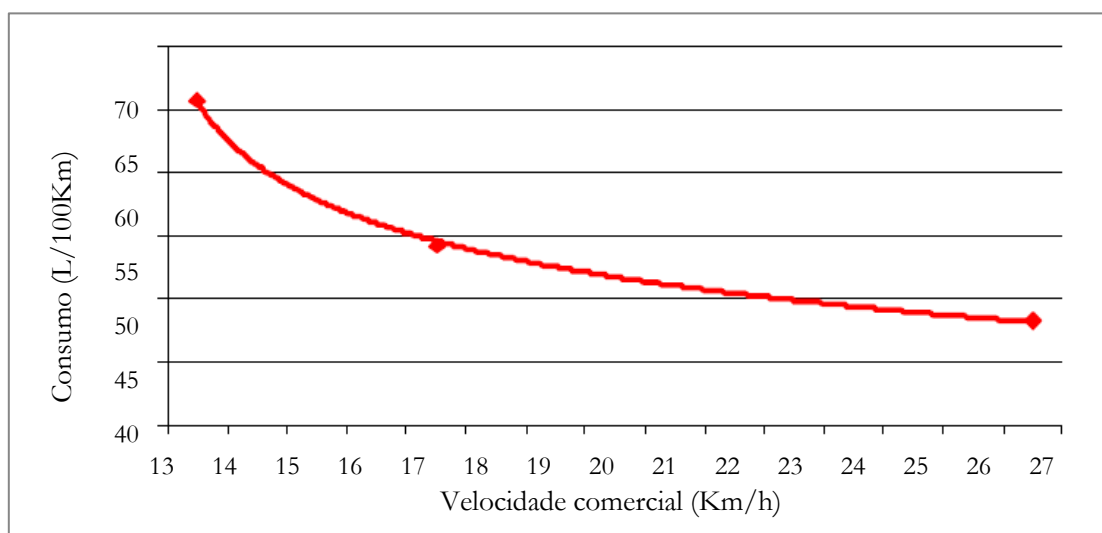


Figura 7 – Consumo (L/100km) vs Velocidade comercial (km/h) (Fonte UITP, 2009)

É de referir ainda que, de acordo com WUNCH, 1996, o impacto da velocidade no custo operacional é muito maior para autocarros do que para os outros modos de transporte público. Estudos recentes no âmbito do processo de reestruturação da Carris estimaram que cada acréscimo de 1

km/h na velocidade comercial permitirá à Carris poupar cerca de 5 milhões de euros, devido às poupanças em pessoal, consumo e necessidades de frota (VIEIRA, 2004).

2.3 A infra-estrutura e os condicionalismos de funcionamento

2.3.1 Caracterização topográfica

A caracterização topográfica de uma rede de autocarros tem sido abordada por vários estudos que comprovam a sua influência na variação do consumo de combustível das viaturas. Em SIMÕES, 2005, para inclinações positivas (sentido da subida) acima de 5%, a correlação com o consumo mostrou-se bastante evidente. Já a correlação entre o consumo e as inclinações negativas não se provou significativa, o que poderá ser explicado pelo facto de a massa do veículo deixar de ser um entrave ao movimento. Também noutros estudos é possível identificar a forte correlação que inclinações elevadas têm com o consumo das viaturas. Segundo KOSHAL, 1970, os custos para as empresas que operam em rotas de montanha são significativamente superiores. HENSHER, 2003, conclui que o terreno tem uma influência significativa no consumo de combustível e no tempo total de viagem de um autocarro em operações não urbanas. LYNCH, 1971, aponta que as zonas planas são mais adequadas à actividade de uma frota. Além disso, os veículos que circulam em zonas de declive mais elevado são mais susceptíveis a avarias, o que resulta em custos operacionais mais elevados (MARSH, 1983). Segundo a VOLVO, 2010, a constante condução em zonas muito declivosas pode aumentar o consumo de combustível em mais de 50%.

A razão para estes resultados pode estar associada ao facto de, em superfícies planas, ser mais fácil manter uma velocidade de circulação constante mas também à menor probabilidade de ocorrência de acelerações bruscas, já que a aceleração brusca súbita injecta mais combustível do que o necessário (PCRA, 2005).

2.4 O motorista

2.4.1 Idade média de motoristas

Alguns estudos já realizados apontam no sentido de que a idade dos motoristas de autocarros tem influência no modo de condução e de reacção dos mesmos. Os condutores de meia-idade (entre os 25 e os 60 anos de idade) apresentam, geralmente, um menor risco de acidente que os condutores jovens (sobretudo do sexo masculino) (WILLIAMS, 1985) e mais velhos (CHIPMAN ET AL., 1993). Quanto aos condutores mais velhos, são também menos propensos a erros de condução (ÅBERG E RIMMÖ, 1998). Em motoristas de veículos pesados, os homens na casa dos 20 anos têm uma taxa de acidentes substancialmente mais elevada (CAMPBELL, 1991).

Os motoristas de autocarro mais jovens circulam a velocidades superiores, mesmo em circunstâncias normais. Ainda assim, com o aumento da pressão temporal do cumprimento de horários, o comportamento e reacções dos motoristas alteram-se, levando o motorista a dirigir a uma velocidade superior. Esta diferença na reacção à pressão é menor para os motoristas mais velhos, já que estes não são tão susceptíveis à pressão do tempo (ZARKADOULA ET AL., 2007).

2.5 A política da empresa de transporte

2.5.1 Eventos de condução

A Rodoviária de Lisboa tem ao seu dispor, em cerca de 1/3 da sua frota, um equipamento (o sistema FM200, do programa GisFrot) que permite detectar e registar eventos de condução que foram considerados como relevantes para a caracterização da condução de um motorista, como as acelerações bruscas, o excesso de velocidade ou a ocorrência de tempo excessivo de ralenti.

Nas tabelas seguintes, encontram-se as evoluções registadas na Rodoviária de Lisboa após a introdução do sistema FM200 em algumas viaturas da frota:

Tabela 1 (Fonte RL, 2008b)

Evolução dos eventos de Conforto			
Eventos	Registados antes de Maio 2006 (segs.)	Registados após Maio de 2006 (segs.)	Variação (%)
Acelerações Bruscas	96	66	-32%
Travagens Bruscas	159	74	-53%

Tabela 2 (Fonte RL, 2008b)

Evolução dos eventos de Segurança			
Eventos	Registados antes de Maio 2006 (segs.)	Registados após Maio de 2006 (segs.)	Variação (%)
Total de travagens	482	496	2,7%
Kick-down ¹	51	3	-95%

¹ Em caixas de transmissão automática, ocorre kick-down para uma mudança mais baixa de forma a adquirir-se mais potência do motor, através do aumento de rotações.

Velocidade excessiva	82	41	-50%
Excesso de rotação	97	68	-30%

Tabela 3 (Fonte RL, 2008b)

Evolução dos eventos de Desempenho ambiental			
Eventos	Registados antes de Maio 2006 (segs)	Registados após Maio de 2006 (segs)	Variação (%)
Excesso de tempo em Ralenti	725 segs	877 segs	21%
Consumo gasóleo	46,72 L/100Km	45,63 L/100Km	-2,3%

Actualmente, já existem várias referências que sugerem a influência destes acontecimentos no consumo de uma viatura. Segundo a VOLVO, 2010, a redução da velocidade de 90 km/h para 80 km/h reduz o consumo de combustível em 6%.

O gráfico seguinte demonstra até que existe uma velocidade óptima de circulação para que as emissões de dióxido de carbono sejam mínimas e, portanto, também o consumo de combustível:

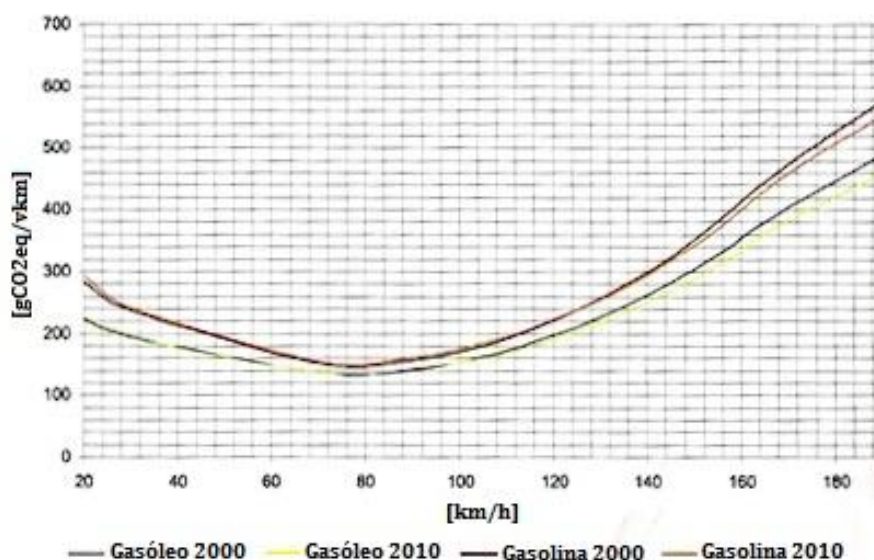


Figura 8 - Variação do factor de emissão de CO2 face à variação da velocidade de circulação. (Fonte: Moura, e tal, 2003)

As acelerações mais bruscas também propiciam um aumento de consumo, como se pode constatar no seguinte gráfico, onde também é evidente o aumento do consumo com a velocidade de circulação.

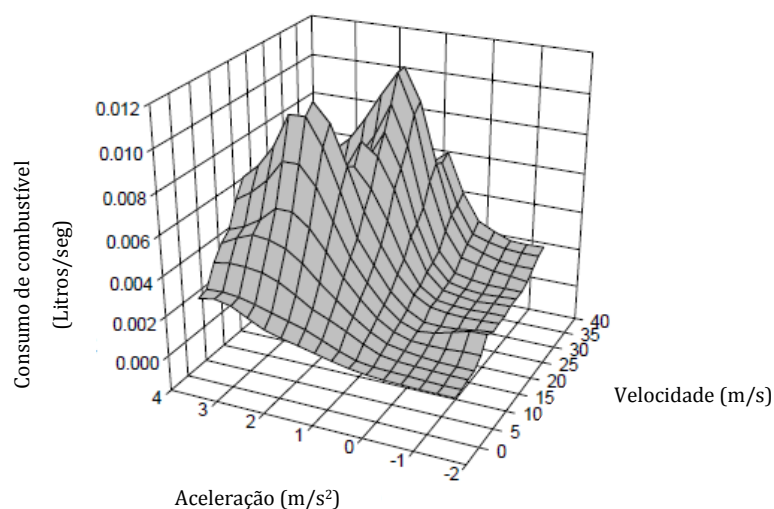


Figura 9 – Consumo de combustível vs Aceleração vs Velocidade (Fonte: KYOUNGHO et al., 2002)

Segundo SIMÕES, 2005, reduções de 10 a 20% na taxa de aceleração poderão significar diminuições entre 2 a 3L/100Km no consumo das viaturas sem levar a perdas nos tempos de viagem.

Outro factor de considerável influência no consumo das viaturas é o tempo em ralenti. Um veículo típico, a gasolina, gasta cerca de 1 litro de combustível por hora ao ralenti e, a gasóleo, cerca de 0,7 litros de combustível por hora ao ralenti (FARIAS, 2009). Também em SIMÕES, 2005, indica-se ainda que uma redução do tempo ao ralenti médio, em estação terminal, de 150 segundos para 60 segundos, corresponderia a uma redução no consumo de combustível até 95000 L por ano (para uma frota de 150 viaturas).

Segundo RAKHA e DING, 2003, as travagens efectuadas por um veículo também influenciam o seu consumo de combustível uma vez que este tipo de eventos provoca aumentos consideráveis tanto nas taxas de emissão de gases como no consumo de combustível dos veículos, especialmente para altas velocidades de cruzeiro.

2.5.2 Número médio de monitorizações formativas de motoristas

Actualmente, a Rodoviária de Lisboa tem um programa de acompanhamento dos seus motoristas em que são realizadas sessões de monitorização formativa junto dos mesmos. Nestas sessões os motoristas são confrontados com comportamentos menos correctos da sua condução, para que esses comportamentos possam ser corrigidos e se obtenham ganhos quer a nível de consumo como a nível de conforto e segurança para o passageiro.

Na Figura 10, encontra-se a evolução do consumo específico da frota da *Rodoviária de Lisboa* após início do programa *GisFrot*.

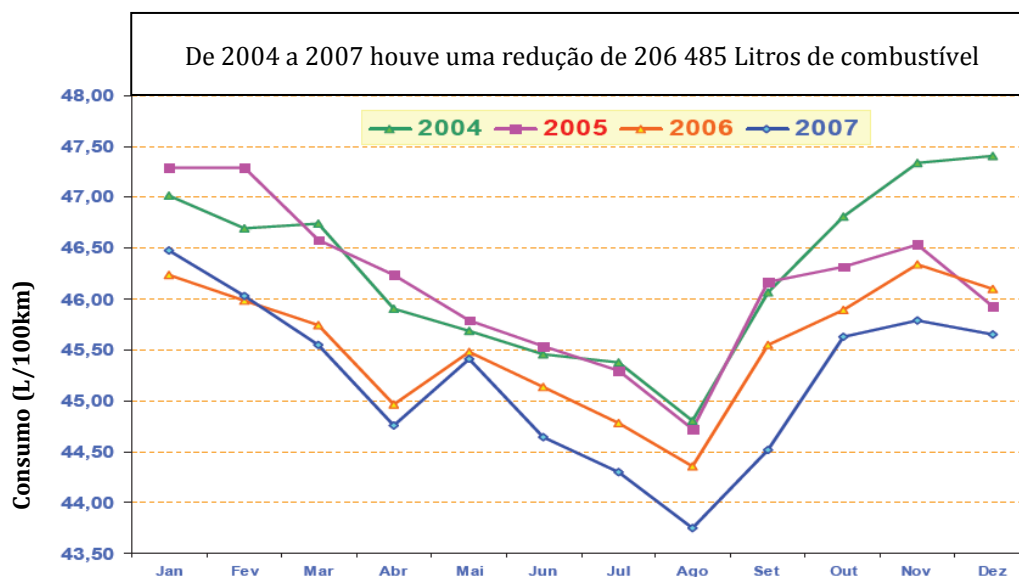


Figura 10 - Evolução do consumo específico de gasóleo (L/100km). (Fonte: RL, 2008)

Pode-se concluir que as monitorizações formativas executadas com os motoristas resultam numa tendência para a diminuição de consumo por parte dos motoristas da RL. Associadas à monitorização dos eventos de condução, as monitorizações formativas, bem como as acções de correcção e de melhoria da sua condução, permitiram que, na *Rodoviária de Lisboa*, se alcançasse uma redução de 2,6% no consumo específico global da empresa, até Abril de 2009. Além deste aspecto, também se deve ter em conta que este processo levará à consolidação de um sentimento de conforto e segurança a bordo por todos os que utilizam os serviços de uma operadora de transporte (RL, 2008a).

O *Centro de Energias Renováveis da Grécia* conduziu um estudo piloto de eco-condução para avaliar os efeitos da mudança do estilo motoristas de autocarros, sendo que os motoristas conseguiram uma média de 10,2% na poupança de combustível durante o período de treino, pelo que foi estimado que seria possível alcançar uma redução de combustível entre 10 a 15% com esse programa (ZARKADOULA ET AL., 2007).

Existem, portanto, indícios de que as iniciativas de educação de condutores de transportes rodoviários tenham potencial de melhoria do desempenho dos mesmos. Dependendo ainda das características da via ou da tecnologia do veículo, têm sido reportadas variações até 40% no consumo de combustível quando comparadas a um condutor agressivo (VLIEGER, 2000).

3. Conceitos estatísticos sobre Regressão Linear

No presente capítulo pretende-se apresentar os conceitos teóricos que suportem as análises estatísticas realizadas aos dados recolhidos e de que forma se irá verificar a validade dos modelos a desenvolver. Inicialmente será feita uma descrição da regressão linear múltipla, como é composta e quais os seus pressupostos estatísticos. Em seguida, serão apresentados os testes a realizar no intuito de verificar e validar os modelos que se pretendem desenvolver no âmbito desta dissertação.

3.1 A regressão linear múltipla

A regressão linear múltipla é um método de análise para avaliar a força da relação entre um conjunto de variáveis explicativas, também designadas como variáveis independentes, e uma única variável de resposta (ou variável dependente).

A regressão linear pode ser utilizada com diversos propósitos, como averiguar a capacidade de um conjunto de variáveis em prever um determinado resultado ou para testar se a adição de uma variável a um modelo contribui para a capacidade preditiva do mesmo. As relações entre variáveis podem ser não lineares e as variáveis independentes podem ser quantitativas ou qualitativas (PALLANT, 2007).

Através de uma análise de regressão múltipla, é obtido um conjunto de resultados conhecidos como coeficientes de regressão, atribuídos a cada variável explicativa. Estes coeficientes exprimem a variação estimada para a variável dependente associada a uma mudança unitária da variável explicativa correspondente, com a condição de que as demais variáveis explicativas permaneçam inalteradas. O ajuste de um modelo de regressão múltipla pode ser julgado de diversas maneiras, quer seja pelo cálculo do coeficiente de ajustamento quer pela análise de resíduos (LANDAU, 2004).

Algebricamente, pode-se adotar um modelo geral com a seguinte forma:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1)$$

Neste modelo, x_1, x_2, \dots, x_k corresponde aos valores das variáveis regressoras, enquanto os parâmetros $\beta_0, \beta_1, \beta_2, \dots, \beta_k$, e ε se denominam por coeficientes de regressão parciais ou não standardizados e expressam a relação entre a variável dependente, y , quando as restantes variáveis independentes, x_1, x_2, \dots, x_k , são mantidas constantes na regressão.

O termo ϵ é o resíduo e representa o desvio entre valor observado da respectiva resposta em relação ao valor esperado obtido pelo modelo. Os resíduos podem ter diversas origens, tais como a ausência de factores influentes no modelo ou até os erros de medição (LANDAU, 2004).

3.2 Coeficientes de regressão não standardizados

Os coeficientes de regressão estimados, denominados coeficientes não standardizados, representam tanto o tipo de relacionamento como a força da relação entre a variável dependente e as variáveis independentes de regressão. O sinal do coeficiente indica se o relacionamento é positivo ou negativo e o valor do coeficiente indica a alteração da variável dependente quando a variável independente é alterada em uma unidade (TABACHNICK E FIDELL, 2007).

3.3 Coeficientes de regressão standardizados

A standardização dos coeficientes de regressão converte-os para uma escala comum, de forma a garantir que os coeficientes de todas as variáveis são comparáveis. Tendo-se uma escala comum para todas as variáveis, pode-se determinar qual a variável que tem o maior impacto na determinação da variável dependente.

Embora os coeficientes standardizados representem uma medida objectiva da importância que desempenha cada variável no modelo, permitindo a sua comparação, devem ser tomadas algumas precauções na sua utilização (TABACHNICK E FIDELL, 2007):

- Só devem ser usados como um guia para a importância relativa de cada variável independente quando a colinearidade for mínima, uma vez que a colinearidade pode distorcer as contribuições de cada variável independente;
- Os valores beta só podem ser interpretados no contexto das outras variáveis presentes no modelo de regressão, já que um modelo com variáveis diferentes implicaria, muito provavelmente, um valor beta diferente para uma mesma variável.

O cálculo dos coeficientes de regressão standardizados, b'_k , é feito através da seguinte expressão (TABACHNICK E FIDELL, 2007):

$$b'_k = \beta_k \frac{s_{x_k}}{s_y} \quad (2)$$

3.4 Métodos de regressão linear múltipla

Existem 3 métodos principais da técnica de regressão múltipla: regressão standard, regressão hierárquica e regressão sequencial, e as diferenças entre estes métodos determinam a ordem de entrada das variáveis independentes na equação.

Regressão standard: Neste método de regressão todas as variáveis independentes são introduzidas, em simultâneo, no modelo. Cada variável independente é então avaliada em termos da variância que consegue explicar (HO, 2006).

Regressão hierárquica: O método de regressão hierárquica é mais flexível que o anterior, uma vez que permite ao investigador determinar a ordem de entrada das variáveis independentes no modelo de regressão. Cada variável independente é avaliada em termos do poder explicativo que acrescenta para o modelo (HO, 2006).

Regressão sequencial: Neste método, a ordem de entrada de variáveis é baseada unicamente em critérios estatísticos. Às variáveis com uma correlação mais forte com a variável dependente é atribuída prioridade de entrada no modelo, sem ter em conta considerações teóricas relativamente ao tema abordado. Inicialmente, este método é utilizado com o intuito de averiguar o poder explicativo das variáveis independentes. A regressão sequencial pode ser feita a partir de 3 métodos de selecção de variáveis: eliminação backward, selecção forward e stepwise (HO, 2006).

3.5 Pressupostos de um modelo de regressão linear

A aplicação de um modelo de regressão linear implica a verificação de alguns pressupostos estatísticos, pelo que é importante verificar a validação dos mesmos ao longo do cálculo dos coeficientes de regressão e previsão da variável dependente. Os pressupostos a serem analisados dividem-se em quatro áreas:

1. Linearidade dos parâmetros;
2. Variância constante dos resíduos - homocedasticidade;
3. Independência dos resíduos;
4. Normalidade da distribuição dos resíduos;
5. Continuidade da variável dependente.

3.5.1 Linearidade

A linearidade da relação entre as variáveis dependente e independentes representa de que forma a variação na variável dependente está associada às variáveis independentes. O coeficiente de regressão é constante em toda a gama de valores da variável independente. O conceito de correlação é baseado numa relação linear e, portanto, torna-se uma questão crucial na análise da regressão. A linearidade de uma relação bivariada é facilmente examinada através de gráficos de resíduos. Qualquer padrão curvilíneo detectado nos gráficos de resíduos indica que a necessidade de uma acção correctiva (como, por exemplo, a transformação das variáveis) com o objectivo de restaurar a linearidade, podendo aumentar tanto a precisão preditiva do modelo como a validade dos coeficientes estimados (TABACHNICK E FIDELL, 2007).

3.5.2 Variância constante dos resíduos - Homocedasticidade

Um dos pressupostos de um modelo de regressão linear é a de que os erros devem ter variância constante. Esta condição é chamada de homocedasticidade. Quando as perturbações são oscilantes, os erros são chamados de heterocedásticos. A verificação da homocedasticidade pode ser feita, entre outros, através da análise gráfica dos resíduos versus valores ajustados, que devem apresentar pontos dispostos aleatoriamente, sem nenhum padrão definido (NETTER ET AL., 1985).

3.5.3 Independência dos resíduos

A independência de resíduos é garantida quando a dimensão de um resíduo não influencia a dimensão do resíduo seguinte. Para isso, é necessário que a correlação entre resíduos sucessivos seja nula. A dependência entre resíduos poderá surgir a partir de observações que sejam sequenciais no tempo (como resultado, por exemplo, o “tempo de retorno” de um aparelho de medição). A melhor forma de identificar se os resíduos são independentes consiste em obter um gráfico que oponha os resíduos aos valores previstos das variáveis. Se os resíduos forem independentes, não deverá ser possível identificar um padrão na sua distribuição (TABACHNICK E FIDELL, 2007).

3.5.4 Normalidade da distribuição dos resíduos

Provavelmente o pressuposto que mais frequentemente é violado é a não normalidade das variáveis dependentes, independentes, ou de ambas. O diagnóstico mais simples para as variáveis

independentes consiste num histograma de resíduos, através da verificação da aproximação da distribuição dos resíduos à distribuição normal. Apesar de ser um método atractivo devido à sua simplicidade, este método é particularmente difícil em amostras de menores dimensões. O melhor método de verificar a normalidade dos resíduos é através do uso de gráficos de probabilidade normal (TABACHNICK E FIDELL, 2007). Além desta hipótese é também possível fazer essa verificação recorrendo a testes paramétricos como os de Kolmogorov-Smirnov e Shapiro-Wilk.

3.6 Multicolinearidade e singularidade

A multicolinearidade refere-se ao relacionamento entre as variáveis independentes. Esta existe quando as variáveis independentes são altamente correlacionadas e leva à inflação do erro padrão dos coeficientes, resultando numa redução da sua importância. Assim, devem ser tomados cuidados na escolha das variáveis independentes de tal forma que elas não sejam altamente correlacionadas entre si.

A singularidade é um fenómeno que ocorre quando uma variável independente é, na verdade, a combinação de outra variável independente submetida ao modelo (GAUR e GAUR, 2009).

3.7 Tolerância e factor de inflação da variância (VIF)

Estes factores são utilizados para fazer o diagnóstico de colinearidade das variáveis do modelo de regressão e são inversos um do outro. Como referência, pode-se tomar valores de VIF superiores a 5, ou uma tolerância inferior a 0.2, como indicadores da presença de multicolinearidade entre variáveis (GAUR e GAUR, 2009).

Apesar de estas estatísticas de diagnóstico serem comuns na detecção de multicolinearidade, a sua eficiência é controversa, pelo que devem ser utilizados com alguma precaução.

Assim que for detectada multicolinearidade entre variáveis no modelo, os coeficientes de regressão devem ser considerados sem significado, pelo que se deve proceder à remoção das variáveis que sejam altamente correlacionadas.

3.8 Dimensão da amostra

A dimensão que deverá ter uma amostra, para que o modelo de regressão tenha significado estatístico, é um assunto sobre o qual vários autores não têm uma posição unânime. A dimensão da amostra tem um impacto directo no ajustamento e no poder de um modelo de regressão linear. Segundo (TABACHNICK E FIDELL, 2007), amostras pequenas, com menos de 30 observações, são apropriadas apenas para análises de regressão linear simples, com apenas uma variável independente. Mesmo nessas situações, só podem ser detectadas relações fortes e com um grau de certeza significativo. Por outro lado, amostras de grande dimensão, como por exemplo com 1000 observações ou mais, tornam os resultados estatísticos demasiado sensíveis, pelo que, geralmente, acabam por indicar que quase todas as relações têm algum significado estatístico. Como regra geral, considera-se que a relação entre o número de observações e variáveis independentes nunca deve ser inferior a 5:1, o que significa que são necessárias 5 observações por cada variável independente. Embora o rácio mínimo seja de 5:1, a relação desejável situa-se entre 15 a 20 observações por cada variável independente. STEVENS, 1996, recomenda cerca de 15 casos por variável preditiva, para que se obtenha uma equação fiável, enquanto TABACHNICK E FIDELL, 2007, fornecem uma fórmula para calcular a dimensão mínima da amostra, tendo em conta o número de variáveis independentes que se pretendem usar no modelo, de $N > 50 + 8v$, em que “v” corresponde ao número de variáveis independentes.

3.9 Outliers

Os *outliers* são observações extremas, não características, que apresentam resíduos que são consideravelmente superiores em valor absoluto, aos resíduos das outras observações. Os efeitos dos *outliers* podem ser moderados se forem encontrados no meio do domínio das observações, ou podem ser extremos caso se encontrem próximos dos limites do domínio das observações. Isto é, o local onde se encontra o *outlier* determina a severidade da sua influência sobre a estimação dos coeficientes de regressão. Para se eliminar os efeitos da magnitude da escala de medida sobre os erros, é comum standardizá-los para que a média se mantenha em zero e o desvio-padrão seja unitário. Existem 2 tipos de resíduos utilizados mais frequentemente na detecção de outliers: resíduos padronizados e resíduos studentizados (GAUR e GAUR, 2009).

3.10 Variáveis *Dummie*

Em modelos onde existam variáveis qualitativas tenham relevância explicativa, podemos recorrer ao uso de variáveis *dummies*. Uma variável *dummie* assume apenas os valores 0 ou 1, para indicar

se determinada condição é satisfeita ou não, para cada indivíduo da amostra. Em geral, quando se tem uma variável qualitativa que representa m atributos diferentes, devem criar-se $m - 1$ variáveis *dummies*. Uma desvantagem da utilização de variáveis *dummies* é o facto de levarem a uma perda de graus de liberdade da regressão, pelo que deve ser avaliada cuidadosamente a adopção deste tipo de variáveis, tendo-se em conta o número total de observações.

A decisão sobre qual a categoria não codificada é muitas vezes arbitrária. A categoria que não é codificada é aquela com a qual todas as outras categorias serão comparadas. Como tal, muitas vezes, a maior categoria será a não codificada.

Um coeficiente de um regressor *dumme* obtido, não standardizado, é interpretável como a diferença esperada entre uma determinada categoria e a categoria não codificada, relativamente à variável dependente. Se um regressor *dumme* for padronizado, esta interpretação deixa de fazer sentido. Este tipo de variáveis só podem ser aplicadas às variáveis independentes (GARAVAGLIA ET AL., 2009).

3.11 Transformação de variáveis

A relação básica entre a variável dependente e as variáveis independentes, representada numa regressão linear múltipla é, como o nome indica, linear. No entanto, muitas vezes as relações entre ambas não são lineares pelo que é necessário proceder a transformações que permitam obter uma maior correlação entre variáveis. As transformações de variáveis servem também como meio de corrigir as violações de alguns pressupostos estatísticos, mas o principal propósito da transformação de variáveis é um dos seguintes:

- Melhorar ou modificar a relação entre variáveis dependentes e independentes;
- Possibilitar a utilização de variáveis não métricas no modelo de regressão.

Vários tipos de transformações de variáveis são apropriados para a linearização de uma relação curvilínea. Abordagens directas envolvem modificação de valores através de uma transformação aritmética, seja através de um logaritmo ou uma raiz quadrada, por exemplo. Ainda assim, as transformações de variáveis têm certas limitações:

- Só são aplicáveis a uma simples relação curvilínea (a relação com apenas um giro ou ponto de inflexão);
- Não representam nenhuma informação estatística de avaliação sobre qual dos modelos, linear ou o curvilíneo, é mais adequado;

- Traduzem-se apenas na relação univariada e não na interacção entre variáveis, quando o modelo tem mais de uma variável independente (TABACHNICK E FIDELL, 2007).

3.12 Metodologias de análise de um modelo de regressão linear

Após a obtenção do modelo de regressão linear múltipla, é de extrema utilidade a criação de gráficos que permitam verificar os pressupostos da regressão já mencionados, bem como de testes que permitam verificar a validade dos pressupostos da regressão linear. Estas análises podem ser agrupadas nas seguintes classes:

- Verificação da linearidade e dos pressupostos de normalidade;
- Detecção de *outliers* e observações influentes;
- Diagnóstico sobre o efeito/influência das variáveis.

3.13 Verificação da adequabilidade da regressão ajustada

A medida de ajuste do modelo é fornecida pelo coeficiente de correlação múltiplo, R , que se define como a correlação entre os valores observados da variável de resposta e os valores previstos pelo modelo. O valor de R^2 exprime o valor da explicação da variável de resposta pelas variáveis independentes, podendo variar entre 0 e 1, e dá-se o nome de coeficiente de determinação. O coeficiente R^2 será nulo no caso de todos os elementos do vector das variáveis explicativas, excepto o termo constante, serem iguais a zero, ou seja, o modelo, não possuir poder explicativo. O coeficiente R^2 terá o valor de 1 caso os resíduos do modelo sejam iguais a zero, ou seja, 100% da variância de y é explicada pelo modelo. Portanto, quanto mais próximo de 1 estiver o valor de R^2 , mais adequado será o modelo de regressão ajustado (CHARNET ET AL., 1999).

A adição de variáveis, implica um aumento do valor de R que, por sua vez, leva ao aumento de R^2 , mesmo que sejam adicionadas variáveis sem significado. O impacto deste efeito é tanto mais notório quanto mais próxima for a dimensão da amostra do número de variáveis independentes. Desta forma, é necessário encontrar uma forma objectiva de medir o ajustamento do modelo sem que este aumente com a inclusão de variáveis sem significado estatístico para o modelo (TABACHNICK E FIDELL, 2007).

O software *SPSS*, que será utilizado para calcular o modelo de regressão, calcula o valor do coeficiente de correlação múltiplo ajustado, $R^2_{ajustado}$. Um valor $R^2_{ajustado}$ de 0.70 significaria que as

variáveis independentes no modelo podem prever 70% da variância da variável dependente (ADA, 2009).

3.14 Análise de variância do modelo de regressão linear

Os cálculos necessários para a obtenção do R^2 podem ser resumidos na Tabela 4, denominada Tabela de análise de variância (ou ANOVA).

Tabela 4 – Tabela de análise de variância (ANOVA) (Fonte: MAROCO, 2007)

Fonte de variação	Somas dos quadrados	Graus de liberdade	Quadrados médios	F
FACTOR (entre as amostras)	$SQF = \sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y})^2$	$k - 1$	$QMF = \frac{SQF}{k - 1}$	$F = \frac{QMF}{QME}$
RESIDUAL (dentro das amostras)	$SQE = \sum_{i=1}^k \sum_{j=1}^{n_j} (\bar{Y}_{ij} - \bar{Y}_i)^2$	$N - k$	$QME = \frac{SQE}{N - k}$	
TOTAL	$SQT = \sum_{i=1}^k \sum_{j=1}^{n_j} (Y_{ij} - \bar{Y})^2$	$N - 1$		

A estatística de teste da ANOVA é dada pela razão entre a variância do Factor, (estimada por divisão da SQF pelos respectivos graus de liberdade: $k-1$) e a variância dos erros (estimada por divisão da SQE pelos respectivos graus de liberdade: $N-k$). Assim, rejeita-se a hipótese nula H_0 se $F \geq_{f1-\alpha, (k-1, N-k)}$ (MAROCO, 2007). A estatística F da ANOVA tem distribuição F-Snedecor quando a variável dependente tem distribuição normal em todas as populações e as variâncias populacionais são homogêneas e serve para testar se nenhuma das variáveis explicativas está relacionada com a variável dependente ou, por outras palavras, que o valor de R^2 é nulo (HO, 2006).

3.15 Verificação dos pressupostos de linearidade e normalidade

Quando o número de variáveis é grande, a tarefa de verificar o pressuposto de linearidade torna-se difícil. No entanto, pode-se examinar as hipóteses de linearidade e normalidade através da análise dos resíduos após o ajustamento de um modelo.

Em seguida, indicam-se as características que impedem a validade de uma regressão linear e que são detectáveis a partir da análise de resíduos (NETER, WASSERMAN e KUTNER, 1985):

- A média dos erros não ser nula;
- Os resíduos não terem variância constante;
- Os termos de erro não serem independentes;
- Os resíduos não serem normalmente distribuídos;
- A função de regressão não ser linear;
- A existência de observações “estranhas” (*outliers*);

Para se proceder a esta análise, pode-se recorrer aos seguintes gráficos:

1. **Gráfico de probabilidade normal de resíduos estandardizados:** Este é um gráfico que representa, no eixo das abcissas, a probabilidade observada acumulada dos erros e, no eixo das ordenadas, a probabilidade acumulada que se observaria se os erros possuísem uma distribuição normal. Se os erros possuírem distribuição normal, então os valores representados neste gráfico devem distribuir-se ao longo de uma recta de declive unitário e que passa na origem (CHATERJEE e HADI, 2006).
2. **Histograma de resíduos:** Este histograma pode ser utilizado para verificar se a variância dos resíduos é normalmente distribuída. Um gráfico em forma de sino, distribuído igualmente em torno de zero indica que a suposição de normalidade é, em princípio, verdadeira. Se o histograma indicar que os resíduos não são normalmente distribuídos, isso sugere que os pressupostos subjacentes ao modelo de regressão podem ter sido violados (CHATERJEE e HADI, 2006).
3. **Gráficos de dispersão de resíduos estandardizados versus valores estandardizados previstos:** Este gráfico permite analisar se a distribuição de resíduos é aleatória em torno de zero (MAROCO, 2007).
4. **Gráficos de dispersão de resíduos studentizados versus valores ajustados (Press) previsto:** Tal como no gráfico anterior, através deste gráfico pretende-se verificar se é cumprido o pressuposto de disposição dos resíduos aleatória em torno de zero, sem se identificar nenhum tipo de padrão. Além disso, este gráfico permite ainda a identificação de *outliers* de forma mais evidente que o gráfico de resíduos estandardizados (MAROCO, 2007).
5. **Gráfico de dispersão valores estandardizados DfFits² versus valores estandardizados previstos:** Este gráfico permite detectar as observações mais influentes na regressão, sendo tanto mais influentes quanto maior, em absoluto, o seu valor de DfFits standardizado (MAROCO, 2007).

² DfFits – Esta medida indica o número de erros-padrão em que o valor de y_j será alterado se a observação multivariada j for removida da análise. Assim, observações com valores de DfFits, em valor absoluto, maior que 2 devem ser consideradas com precaução na regressão linear.

3.16 Testes paramétricos

Para testar a significância de factores capazes de influenciar a resposta da variável de medida, a comparação de parâmetros populacionais é um tipo de inferência estatística muito útil. Para isso, podem utilizar-se testes paramétricos que exigem que a forma da distribuição amostral seja conhecida (MAROCO, 2007).

Um dos testes paramétricos mais utilizados é o de Kolmogorov-Smirnov e permite testar a normalidade de uma distribuição. Este exige a verificação simultânea de duas condições: que a variável dependente possua distribuição normal e que as variâncias populacionais sejam homogêneas no caso de se estarem a comparar duas ou mais populações.

3.16.1 Teste Kolmogorov-Smirnov

Utiliza-se então o teste de Kolmogorov-Smirnov para averiguar se a distribuição da variável em estudo ($F(X)$) numa determinada amostra provém de uma população com uma distribuição específica $F_0(X)$ que, neste caso, é a distribuição normal, com parâmetros μ e σ quaisquer que estes sejam. Pretende-se assim testar:

$$H_0: X \sim N(\mu, \sigma) \text{ vs. } H_1: X \sim \neq(\mu, \sigma)$$

Para determinar a estatística de teste é preciso calcular a diferença acumulada que essa observação teria se a sua distribuição de probabilidade fosse normal, bem como a mesma diferença relativa à observação. O valor crítico da distribuição da estatística encontra-se tabelado e rejeita-se H_0 se $D \geq D_{\text{tabela}}(\alpha)$. Assim, rejeita-se H_0 se $p \leq \alpha$. O p-value produzido pelo SPSS é calculado usando a aproximação analítica da estatística de teste de Lilliefors proposta por Dallal & Wilkinson (MAROCO, 2007).

3.16.2 Teste Shapiro-Wilk

Da mesma forma, é possível utilizar o teste Shapiro-Wilk, que possui uma estatística de teste denominada W . Valores pequenos de W sugerem que a distribuição da variável em estudo não é de tipo Normal e os valores críticos para W estão tabelados (PEARSON E HARTLEY, 1972). Este teste é particularmente apropriado, e até preferível ao teste de Kolmogorov-Smirnov, para amostras de pequena dimensão ($N < 30$) (MAROCO, 2007).

3.17 Estimação de elasticidades

É muito frequente estimar-se a capacidade de resposta e a sensibilidade de uma variável dependente Y relativamente às alterações em uma ou em várias variáveis independentes X. Embora os coeficientes de regressão contêmam esta informação, por vezes é mais interessante expressar essa sensibilidade em termos de percentagem da variação da variável dependente aquando de uma alteração de 1% da variável independente. Este é o conceito de elasticidade e é útil na medida em que é adimensional, contrariamente aos coeficientes de regressão estimada, que dependem das unidades da variável.

A elasticidade da variável independente X em relação a uma variável dependente Y é dada pela seguinte expressão:

$$e_i = \beta_i \frac{X_i}{Y_i} \approx \frac{\partial Y_i}{\partial X_i} \times \frac{X_i}{Y_i} \quad (3)$$

A Tabela 5 sumariza a estimação de elasticidade para quatro tipos de transformações habitualmente utilizadas em modelos de regressão (WASHINGTON ET AL., 2003).

Tabela 5 - Sumário de elasticidades face ao tipo de transformação das variáveis (Fonte: WASHINGTON ET AL., 2003)

Modelo	Elasticidade
Linear	$\frac{\Delta Y_i}{\Delta X_i} \times \frac{X_i}{Y_i} = \beta_i \left(\frac{X_i}{Y_i} \right)$
Log-Linear	$\frac{\Delta Y_i}{\Delta X_i} \times \frac{X_i}{Y_i} = \beta_i (X_i)$
Linear-Log	$\frac{\Delta Y_i}{\Delta X_i} \times \frac{X_i}{Y_i} = \beta_i \left(\frac{1}{Y_i} \right)$
Log-Log	$\frac{\Delta Y_i}{\Delta X_i} \times \frac{X_i}{Y_i} = \beta_i$

4. Descrição das amostras estatísticas e desenvolvimento das variáveis

O conjunto de dados recolhido pode ser dividido em três grupos, correspondentes aos vectores energéticos que já foram indicados. No presente capítulo, irá proceder-se a uma descrição e apresentação geral das amostras estatísticas recolhidas e, ainda, à apresentação das variáveis introduzidas no modelo e à metodologia utilizada para o seu cálculo. Pretende-se, ainda, explicar de que forma se espera que cada variável se relacione com o consumo de combustível. No fim, são referidas ainda as limitações detectadas na recolha e utilização de dados.

A recolha de dados cingiu-se apenas aos meses de Outubro de 2009, Novembro de 2009 e Março de 2010, numa primeira fase, mas também Maio de 2010 e Setembro de 2010, tendo estes sido utilizados meramente para validação dos resultados obtidos, uma vez que foram considerados meses padrão ao nível dos movimentos rodoviários, já que não há grandes interrupções no calendário escolar ou laboral nesses períodos. Os dados recolhidos relativos ao movimento das viaturas nas respectivas carreiras, com os quais se obtiveram os resultados dos modelos, correspondem, assim, a um período temporal de 92 dias.

Os dados recolhidos no âmbito desta dissertação, podem dividir-se em 5 áreas principais, desde o tipo de veículo, a caracterização física de uma carreira ao nível das suas infra-estruturas e condicionamentos de funcionamento, mas também o tipo de condução efectuada pelos motoristas e a política da empresa. Em seguida será descrito como se obtiveram as variáveis recolhidas para cada um destes grupos.

4.1 Descrição da amostra de Carreiras

Na recolha da amostra para desenvolvimento do modelo de carreiras foram recolhidos dados relativos a 87 carreiras. Na primeira recolha foram postas de parte as carreiras pertencentes ao projecto Rodinhas, uma vez que não era possível caracterizá-las da mesma forma que as outras, nomeadamente, em relação à caracterização do seu percurso, visto que as paragens não são definidas da mesma forma.

4.1.1 Consumo

Como se pode observar através da Figura 11, a variação de consumo registada para as diferentes carreiras é bastante acentuada, sendo os menores de 20 L/100Km e os maiores consumos na ordem dos 60 L/100Km. É de esperar que as maiores variações estejam directamente relacionadas com o tipo de veículo predominante numa carreira, conclusões que se esperam retirar a partir dos modelos desenvolvidos.

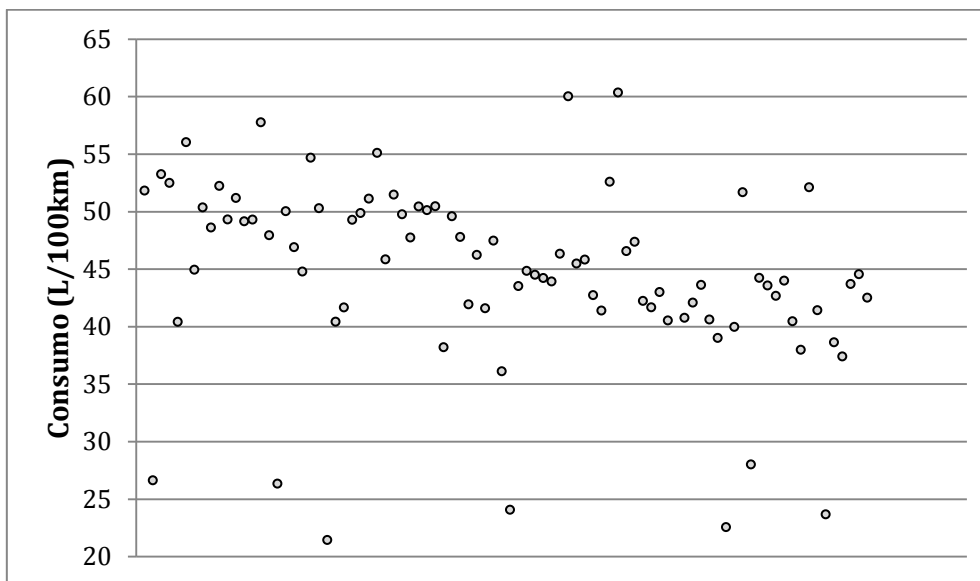


Figura 11 - Consumo de combustível médio para a amostra de carreiras recolhida.

4.1.2 Tipologia de carreira

Através da observação da Figura 12, é possível concluir que a tipologia de carreiras dominante na Rodoviária de Lisboa é a suburbana. Nos modelos a desenvolver será interessante averiguar se a tipologia de uma carreira poderá estar relacionada com o seu consumo, apesar de não existir muita variedade relativamente a carreiras directas e urbanas, o que poderá dificultar a obtenção dessas conclusões.

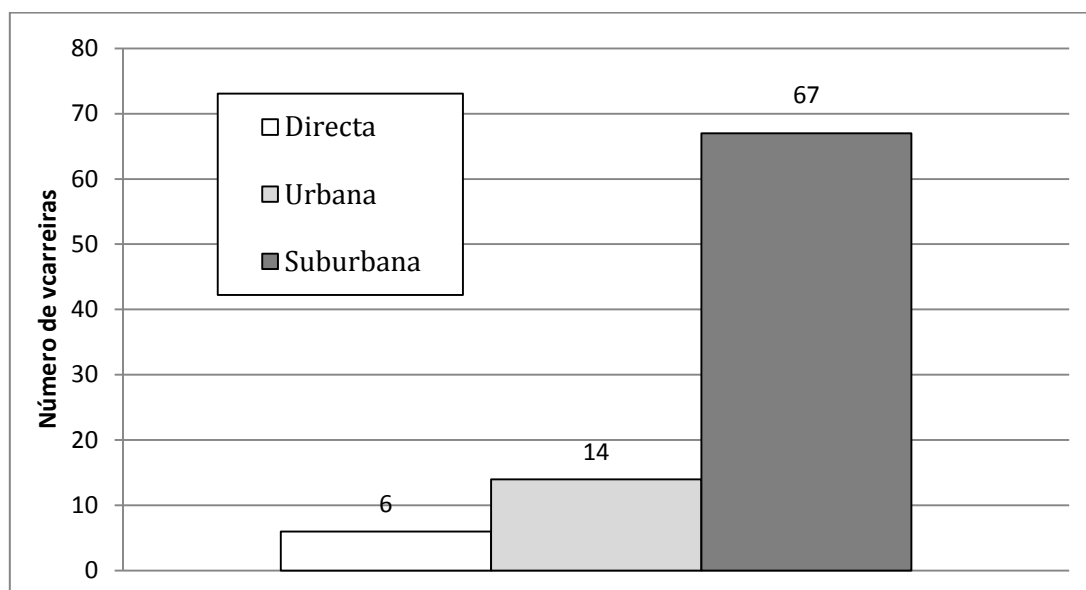


Figura 12 - Número de carreiras por tipologia para a amostra recolhida.

4.1.3 Inclinação

Para se caracterizar a inclinação de cada carreira foram criados vários intervalos de inclinação. O seguinte gráfico de barras indica três intervalos de percentagem de inclinação para as 87 carreiras seleccionadas, obtendo-se um panorama geral sobre as percentagens de cada categoria de inclinação em todos os casos analisados.

Também se desenvolveu uma variável com um intervalo de inclinação superior a 8%, no entanto os seus valores não são apresentados na Figura 13, que se segue, porque se considerou o exemplo apresentado mais relevante.

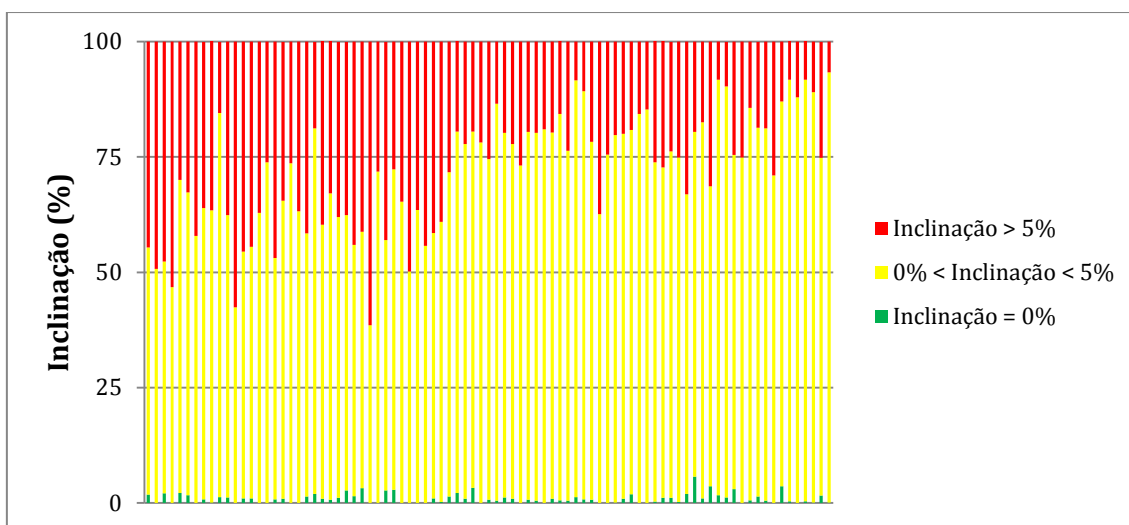


Figura 13 - Percentagem de inclinação para diferentes patamares de inclinação da amostra recolhida.

Na Tabela A1, em anexo, apresentam-se os principais parâmetros estatísticos de caracterização das variáveis relativas às carreiras.

4.2 Descrição da amostra de Motoristas

A recolha de informação relativa a motoristas foi aquela em que se obteve a maior amostra, com um total de 510 motoristas. Ainda assim, foi necessário fazer alguma filtragem a esta amostra, uma vez que foram detectados motoristas que tinham efectuado poucos quilómetros de serviço no período seleccionado, o que se provou ser prejudicial aos dados recolhidos através da análise dos mesmos. Desta forma, foram removidos da amostra os motoristas com menos de 300 kms percorridos por mês, tendo-se eliminado assim 5 motoristas da amostra inicial. Além disso, também foi necessário fazer alguma filtragem de motoristas que apresentavam valores de consumo de combustível muito afastados dos valores espectáveis, tendo-se removido aqueles que apresentavam consumos mais díspares relativamente às tipologias conduzidas, pois essas diferenças podem ser justificadas por erros de recolha de dados, tendo em conta limitações que serão indicadas posteriormente. No total, restaram 488 motoristas para introduzir no modelo, o que corresponde a um corte de apenas 4,5% do universo de motoristas recolhido.

4.2.1 Consumo

A Figura 14, relativa aos consumos registados pelos motoristas, demonstra resultados próximos ao verificado para os consumos de carreiras, visto que a recolha de dados foi feita no mesmo período. No entanto, é possível constatar que os valores extremos são mais evidenciados, o que indicará que há condutores a fazer percursos predominantemente com viaturas articuladas e outros com viaturas minis.

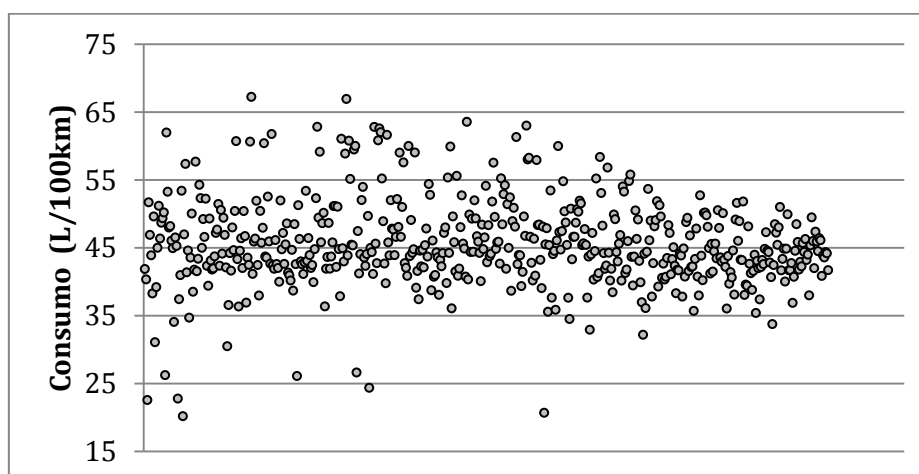


Figura 14 - Consumo de combustível médio para a amostra de motoristas recolhida.

4.2.2 Idade de motoristas

A idade dos motoristas também é um parâmetro que foi detectado, em estudos anteriores, como sendo influente na previsão do consumo de combustível. Desta forma, será interessante verificar se a amostra recolhida na RL permitirá ter uma variabilidade suficiente para que os resultados permitam tirar as conclusões esperadas. Através da análise da Figura 15, relativa à idade de motoristas, pode-se concluir que a amostra apresenta valores muito variados de idade pelo que estarão reunidas as condições para verificar, ou não, os resultados de outros testes desenvolvidos relativamente à sua influência no consumo de combustível médio de um motorista.

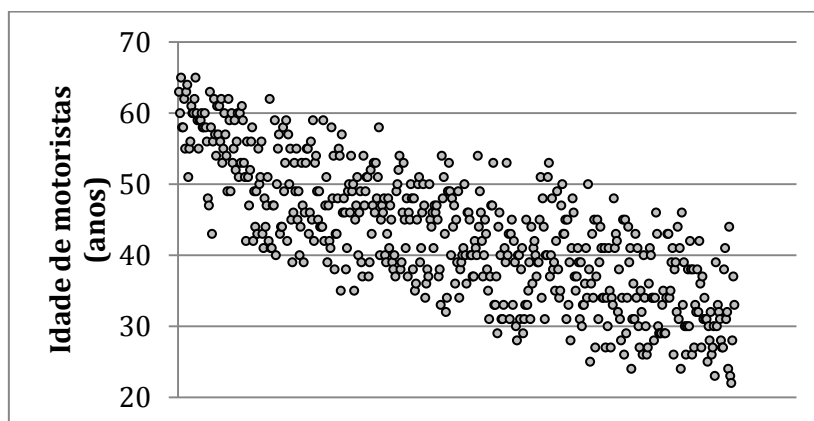


Figura 15 - Idade de motoristas para a amostra recolhida.

4.2.3 Número de monitorizações formativas

O número de sessões de monitorização efectuadas com cada motorista é uma variável que poderá permitir perceber até que ponto este tipo de acções desenvolvidas pela *Rodoviária de Lisboa* junto dos seus trabalhadores terão os efeitos esperados, quer ao nível da redução da percentagem de ocorrência de eventos por parte de cada um dos motoristas, análise que já é actualmente efectuada, mas também para verificar se as alterações na condução dos motoristas levam a uma diminuição do consumo de combustível. Na Figura 16, apresenta-se a variabilidade relativa a esta variável nos motoristas da RL

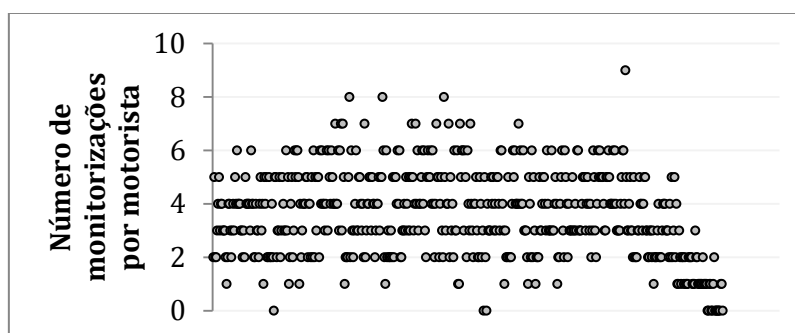


Figura 16 - Número de monitorizações por motorista para a amostra recolhida.

Na Tabela A2, em anexo, apresentam-se os principais parâmetros estatísticos de caracterização das variáveis relativas aos motoristas.

4.3 Descrição da amostra de Viaturas

Na recolha de dados relativos às viaturas, apenas se conseguiu utilizar, para efeitos de criação de um modelo de regressão linear, 105 viaturas. Apesar do número de viaturas ser muito superior (na ordem das 400 viaturas), foi necessário eliminar bastantes observações, visto que só cerca de um terço da frota da RL se encontra equipada com o equipamento FM200, o que impossibilita a obtenção da informação relativa aos eventos de condução, sobre os quais existe um grande interesse ao nível da sua relevância no consumo de combustível das viaturas. Desta forma, a dimensão da amostra é apenas ligeiramente superior à obtida para o modelo de carreiras.

4.3.1 Consumo

Através da Figura 17, apresentada em seguida, é possível ter uma boa noção da distribuição das médias de consumo registadas por cada viatura pertencente à amostra recolhida. Em comparação com os mesmos gráficos obtidos anteriormente para carreiras e motoristas, é notório que os extremos são mais afastados, estando os valores mínimos de consumo médio abaixo dos 20L/100Km e o máximo muito próximo dos 80L/100Km.

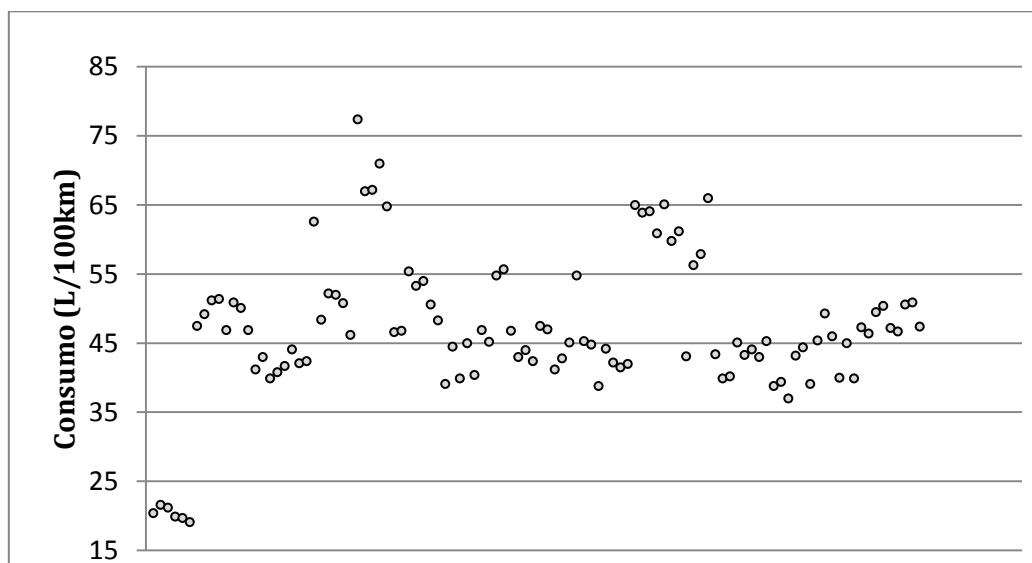


Figura 17 - Consumo de combustível de viaturas para a amostra recolhida.

4.3.2 Tipologia de viaturas

No gráfico de barras que em seguida se apresenta, demonstra-se que não se conseguiu obter nesta amostra indivíduos representantes das 4 tipologias de viaturas. Tal como já foi referido anteriormente, apenas um terço das viaturas se encontra equipada com o equipamento FM200 e, até à data da recolha de dados, nenhuma viatura do tipo Midi se encontrava equipada, pelo que não têm qualquer influência na definição do modelo de regressão de viaturas.

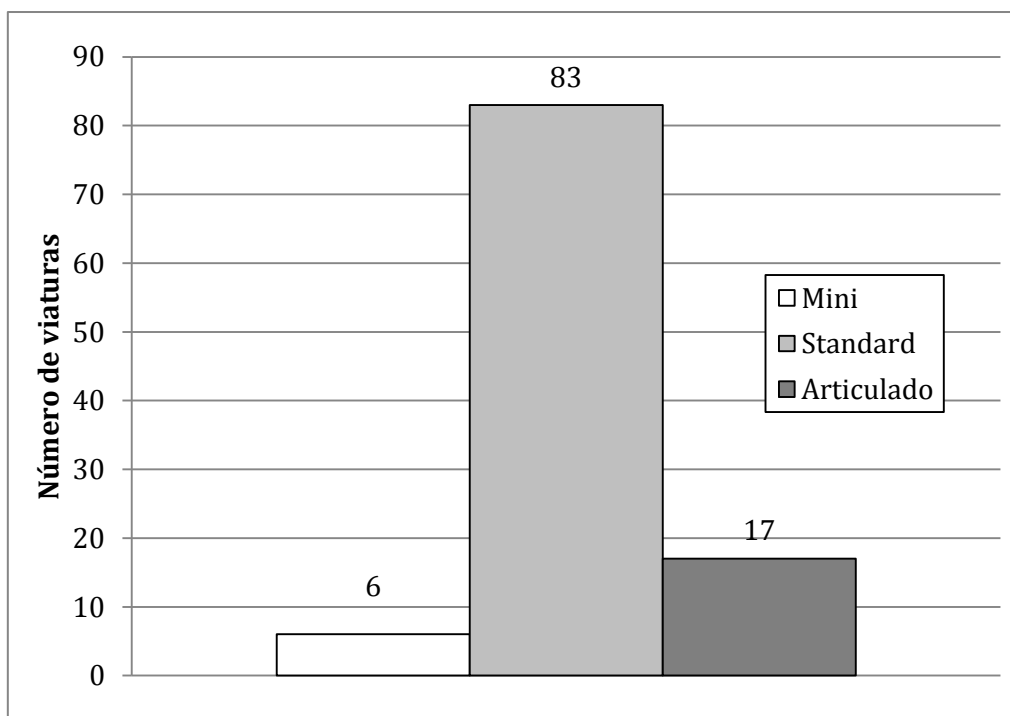


Figura 18 - Número de viaturas por tipologia para a amostra recolhida.

4.3.3 Idade de viaturas

Outro parâmetro que se pretende averiguar se terá influência nos consumos de combustível das viaturas é a sua idade. Através da Figura 19, é possível constatar que a idade dos vários veículos disponíveis na *Rodoviária de Lisboa* apresenta uma grande variabilidade.

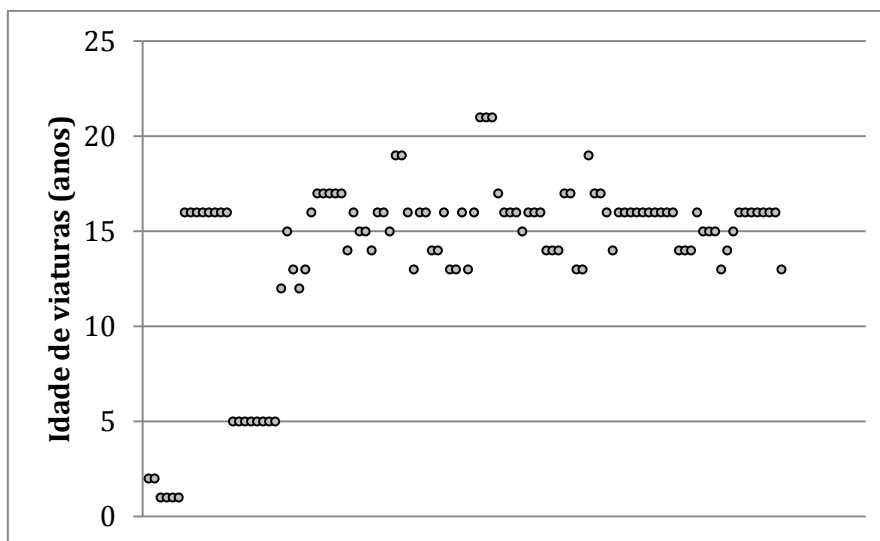


Figura 19 - Idade de viaturas para a amostra recolhida.

Fazendo uma análise mais detalhada desta variável, deve-se salientar que as viaturas mais recentes (com menos de 5 anos) pertencem todas à tipologia Mini, sendo que existem ainda algumas de tipologia Standard com precisamente 5 anos. Relativamente à tipologia de Articulados, todas as viaturas apresentam mais de 10 anos de idade, sendo que a média de idades deste tipo de viaturas é de 15,6 anos. Isto demonstra que existe uma política por parte da RL em abdicar das viaturas mais pesadas, como são os articulados, de acordo com o projecto de renovação de frota da empresa, não só por terem consumos de combustível superiores mas também porque apresentam mais problemas de manutenção relativamente a veículos mais leves e mais modernos.

Na Tabela A3, em anexo, apresentam-se os principais parâmetros estatísticos de caracterização das variáveis relativas às viaturas.

4.4 Descrição dos processos de criação de variáveis

Para se introduzir os dados recolhidos nos modelos de regressão, é necessário proceder-se à transformação das mesmas. Assim, uma vez que se optou por fazer um modelo em função de carreiras, outro sobre motoristas e outro sobre viaturas foi necessário transformar as variáveis para que se pudessem associar os seus valores ao respectivo sujeito de cada modelo. Para se conseguir relacionar os vários tipos de dados, foram utilizados o software *Excel*, para cálculo de variáveis de menor dimensão, como inclinações e distâncias entre paragens, e o software de gestão de bases de dados *MySQL*, com o qual foi possível relacionar, com cada carreira, motorista e viatura, as suas velocidades comerciais, o tempo médio de viagem, as percentagens por quilómetro de cada evento, o número de monitorizações formativas médio de motorista, o número médio de anos de

trabalho de motorista na RL, os dados relativos ao consumo médio dos veículos, a tipologia de veículos utilizada e a idade média de veículos (no Anexo IV apresenta-se o código desenvolvido em *SQL* para este efeito).

Os dados recolhidos no âmbito desta dissertação, podem ser divididos em 5 áreas principais, desde o tipo de veículo, a caracterização física de uma carreira ao nível das suas infra-estruturas e condicionamentos de funcionamento, mas também o tipo de condução efectuada pelos motoristas e a política da empresa.

Em seguida indica-se, detalhadamente, de que forma foram criadas as variáveis inseridas nos modelos de regressão.

4.4.1. Tipo de veículo

4.4.1.1 Tipologia do veículo

Para introduzir no modelo esta informação, é necessário produzir pelo menos 3 variáveis aleatórias independentes, uma vez que se consideraram 4 tipologias distintas de veículos (Minibus, Midi, Standard ou Articulado). De qualquer forma, foram criadas as variáveis correspondentes a cada tipologia para que, aquando da definição do modelo, fosse possível averiguar qual, ou quais, as mais adequadas ao modelo de regressão. Para isso, recorreu-se à seguinte expressão, de forma a adaptar as variáveis aos modelos de Carreiras e Motoristas:

$$PercentagemVeículoTipoY = \frac{\sum QuilómetrosEmVeículoTipoY}{NúmeroTotalQuilómetros} \times 100\% \quad (4)$$

Deve referir-se ainda que, no modelo de viaturas, foram utilizadas variáveis *dummy* para descrever o tipo de viatura em questão.

Estas variáveis são, à partida, de uma importância crucial na definição dos modelos, visto que, de acordo com a tipologia de veículo utilizada, é de esperar que o consumo verificado se encontre entre determinados intervalos de valores distintos para cada tipologia.

Dessa forma, com estas variáveis espera-se que os modelos consigam, interpretar a variabilidade que existe para cada tipologia de viatura.

4.4.1.2 Massa de veículos

A massa de veículos foi introduzida nos modelos de carreiras e de motoristas como a sendo a massa média das viaturas destacadas para uma determinada carreira e para um determinado motorista. A

metodologia utilizada para calcular a massa média das viaturas, por carreira ou por motorista, foi a seguinte:

$$MassaMédiaVeículos = \frac{\sum MassaVeículo}{NúmeroTotalViagens} \quad (5)$$

Por outro lado, no modelo de viaturas, cada uma apresenta o seu próprio valor de massa, de acordo com a tipologia de veículo a que pertença.

O interesse desta variável é o mesmo das variáveis de tipologia de veículo descritas anteriormente. A sua utilidade prende-se com o facto de se poder, com apenas uma variável, descrever o mesmo que se descreve com um mínimo de 3 variáveis distintas de tipologia de veículo. No modelo relativo a viaturas, é de esperar que a variável relativa à tipologia de veículo seja definida apenas pela variável de massa.

4.4.1.3 Idade média de veículos

Esta variável foi calculada de forma idêntica às variáveis de número médio de monitorizações e de anos de trabalho de motoristas, através da seguinte expressão para os modelos de Carreiras e Motoristas:

$$IdadeMédiaVeículos = \frac{\sum IdadeVeículo}{NúmeroTotalViagens} \quad (6)$$

No modelo de viaturas o valor adoptado em cada caso é o valor preciso da idade de cada viatura.

Da mesma forma que se pretende verificar se a idade/experiência dos motoristas terá influência no consumo de combustível, seria interessante aferir se o desenvolvimento das viaturas ao longo dos anos é suficiente para indicar uma melhoria na sua eficiência energética. Apesar de, hoje em dia, se desenvolverem motores mais eficientes, a introdução de novos elementos nas viaturas, nomeadamente os sistemas de ar condicionado, pode levar a que a melhoria na eficiência do deslocamento não seja suficiente para cobrir estes gastos adicionais.

4.4.1.4 Consumo médio

Na Rodoviária de Lisboa, a obtenção dos consumos de combustível para cada viatura é diária, independentemente da sua utilização numa ou em várias carreiras. Para se obter o consumo médio por Carreira ou Motorista (em L/100km), de forma a introduzir-se esse valor no modelo de

regressão, é necessário recorrer a uma média ponderada de consumo a partir dos quilómetros percorridos pelo veículo, em cada carreira ou por cada motorista:

$$Consumo_{médio} = \frac{\sum_i Consumo_{Dia-i} \times QuilómetrosPercorridos_{Dia-i}}{\sum QuilómetrosPercorridos} \quad (7)$$

4.4.2. Caracterização da carreira

4.4.2.1 Tipologia de carreiras

A variável relativa à tipologia de carreiras foi criada como a percentagem de quilómetros que cada motorista ou cada viatura tenham percorrido, em cada tipologia de carreira, durante o período de recolha de dados:

$$PorcentagemCarreiraTipoY = \frac{\sum QuilómetrosEmCarreiraTipoY}{NúmeroTotalQuilómetros} \times 100\% \quad (8)$$

Já para o modelo de carreiras foram utilizadas variáveis *dummy* que descrevam o tipo de carreira, já que não faria sentido utilizar esse tipo de variáveis descritas para os outros 2 modelos.

Pretende-se com a inclusão desta variável nos modelos saber se são detectáveis diferenças de consumo de combustível evidentes para as 3 tipologias de carreiras analisadas: Urbana, Suburbana e Directa. Uma vez que as características do percurso de cada um destes tipos de carreira são consideravelmente distintas, nomeadamente na distância entre paragens e também no comprimento do próprio percurso e, portanto, nas velocidades comerciais de cada um, poderão verificar-se diferenças consideráveis nos seus consumos de combustível.

4.4.2.2 Comprimento de percurso, distância entre paragens e número de paragens

Para caracterizar as carreiras foram desenvolvidas as variáveis comprimento de percurso, número de paragens e distância entre paragens (média, mínima e máxima). Naturalmente, para o modelo de carreiras, será atribuído o valor preciso de cada uma destas variáveis de caracterização de cada carreira. Já nos modelos de motoristas e viaturas será utilizada a média verificada para todos os percursos efectuados relativamente ao número total de viagens realizadas. Desta forma, apenas se explicita a metodologia de cálculo de uma das variáveis, visto que a das restantes é em tudo semelhante:

$$\text{ComprimentoPercurso}_{\text{médio}} = \frac{\sum \text{ComprimentoCarreira}_i}{\text{NúmeroTotalViagens}} \quad (9)$$

O interesse destas variáveis prende-se com a verificação da sua influência no desempenho energético aquando da realização dos percursos definidos para cada carreira. A título de exemplo, pretende-se averiguar se, em carreiras com mais paragens ou com paragens muito próximas umas das outras, as perdas de eficiência ao nível do consumo são identificadas como sendo provocadas por características menos indicadas da localização ou número das paragens de recolha de passageiros.

4.4.2.3 Duração de viagem média

A duração de viagem média foi calculada a partir da média de todas as observações de duração de viagem, obtidas pela diferença entre a hora de chegada e a hora de partida de cada carreira, em ambos os sentidos da mesma. Embora os tempos de viagem em cada sentido possam ser bastante diferentes e tendo ainda em conta os horários de ponta, considerou-se que, com o número de observações registado, se obteria um valor razoável de tempo médio de viagem mesmo sem diferenciar o sentido das mesmas, já que este aspecto iria criar problemas ao nível da definição de novas variáveis e de que forma estas se integrariam no modelo.

$$\text{DuraçãoViagem}_{\text{média}} = \frac{\sum \text{DuraçãoViagens}}{\text{NúmeroTotalViagens}} \quad (10)$$

O interesse desta variável prende-se com uma análise que pode ser feita relativamente à própria tipologia da carreira, visto que as durações de percurso de carreiras urbanas e suburbanas deverão ser distintas. Serve ainda para verificar se a duração terá uma influência detectável no consumo de combustível por quilómetro, já que para durações mais pequenas poderão estar associados estilos de condução menos indicados a um melhor desempenho energético de uma viatura.

4.4.2.4 Velocidade comercial média

O cálculo da velocidade comercial média de cada carreira foi obtido através da média das observações de velocidade comercial obtidas em cada viagem, para o respectivo sujeito, seja ele a carreira, o motorista ou a viatura. Desta forma, como a dimensão da amostra é relativamente grande (recolha de dados corresponde a 3 meses de actividade dos vários intervenientes na RL), considera-se que o valor obtido será muito próximo daquele que mais vezes se verifica para ambos os sentidos. A expressão que permite calcular esta variável é então a seguinte:

$$VelocidadeComercial = \frac{\Delta ComprimentoPercurso(km)}{\Delta TempoViagem(h)} \quad (11)$$

O interesse da velocidade comercial é, em parte, semelhante ao da variável de duração de percurso. Ainda assim, sabe-se que um dos principais problemas apontados a transportes colectivos será a sua baixa velocidade comercial o que, em princípio, poderá levar também a baixos valores de eficiência energética das viaturas. Através desta variável espera-se verificar se o consumo de uma viatura é realmente afectado de forma perceptível.

4.4.3. A infra-estrutura e condicionalismos de funcionamento

4.4.3.1 Inclinação

A caracterização topográfica de cada carreira foi considerada, desde logo, como uma variável imprescindível na definição de um modelo de previsão de consumo de combustível. Uma vez que não existia, na Rodoviária de Lisboa SA, qualquer informação recolhida sobre este tema, foi necessário criar todos os dados necessários ao modelo. A RL disponibilizou, em formato digital, os percursos de cada carreira. A estes percursos era, então, necessário atribuir a informação altimétrica de cada ponto. Para isso recorreu-se ao sítio "<http://www.gpsvisualizer.com/elevation>" onde foram inseridas as coordenadas do percurso de cada carreira e se obteve a sua altimetria.

Desta forma, o percurso de cada carreira ficou caracterizado nas três dimensões espaciais. Para se proceder ao cálculo da inclinação entre cada troço, obteve-se, através do software *Excel*, a variação entre distâncias no plano e entre valores altimétricos consecutivos. A inclinação foi determinada através da seguinte expressão:

$$Inclinação = \frac{VariaçãoCotasAltimétricas}{DistânciaEntrePontosNoPlano} \times 100\% \quad (12)$$

Para se transformar esta informação em variáveis utilizáveis num modelo de regressão linear, foram definidas classes de inclinação (0%, inferior a 5%, superior a 5% e superior a 8%) às quais foi atribuída a percentagem de quilómetros efectuados por cada sujeito. Nos modelos de Motoristas e Viaturas associou-se a cada variável a média ponderada, através do número de quilómetros percorridos por estes intervenientes.

Pretende-se com estas variáveis confirmar-se se, para percursos com inclinações superiores, o consumo verificado tende também a ser maior.

4.4.4. O motorista

4.4.4.1 Número médio de anos de trabalho de motoristas

A variável de número médio de anos de trabalho é definida pela seguinte expressão:

$$NúmeroAnosTrabalho_{médio} = \frac{\sum NúmeroMédioAnosTrabalho}{NúmeroTotalViagens} \quad (13)$$

Com a introdução desta variável nos modelos, pretende-se verificar se a experiência dos motoristas terá influência directa nos consumos de combustível verificados. Se à partida seria de esperar que com o aumento dos anos de experiência, o desempenho de um motorista fosse melhorando, por outro lado, existem estudos que indicam que, para motoristas com mais anos de experiência, os níveis de desempenho acabam por sugerir um maior desleixo por parte dos mesmos, apresentando consumos de combustível superiores bem como maior incidência de práticas de condução erradas.

4.4.4.2 Idade média de motoristas

A metodologia para calcular esta variável é em tudo idêntica à anterior, tendo-se utilizado a seguinte expressão:

$$IdadeMotoristas_{média} = \frac{\sum IdadeMotoristas}{NúmeroTotalViagens} \quad (14)$$

Também a justificação da utilização desta variável é semelhante à que se fez anteriormente para o número de anos de trabalho por motorista. No entanto, optou-se por incluir ambas as variáveis nos modelos para verificar qual das duas apresenta uma relação mais perceptível com o consumo de combustível verificado.

4.4.4.3 Eventos de condução

Uma vez que muitas viaturas da RL estão equipadas com o sistema FM200, que possui uma caixa negra para registar certos tipos de eventos definidos, decidiu-se tirar partido destas condições para tornar estes eventos em variáveis, avaliando a sua influência nos consumos de combustível para cada um dos modelos a desenvolver. As variáveis relativas a cada evento foram todas criadas através da seguinte fórmula:

$$PercentagemEventoY_{médio} = \frac{NúmeroOcorrênciasEventoY}{NúmeroTotalEventos} \times 100\% \quad (15)$$

Em seguida, procede-se à descrição de cada evento:

- Evento 0100: Agrega as situações de desconforto resultantes da aceleração longitudinal brusca. É criado a partir da soma dos eventos 1005 e 1070.
- Evento 0110: Agrega as situações de desconforto resultantes da desaceleração longitudinal brusca. É criado a partir da soma dos eventos 1007 e 1015.
- Evento 0120: Identifica o carregamento do ar com a viatura ao ralenti. É criado a partir do evento 1009 sempre que a viatura não esteja em andamento.
- Evento 0121: Identifica o motor em funcionamento ao ralenti, associado a períodos de tempo entre duas viagens. É criado a partir do evento 1009, sempre que a condição do evento 0120 não se verifique.
- Evento 0130: Identifica velocidades excessivas para o serviço que a viatura está a efectuar.
- Evento 0140: Identifica o carregamento do ar em excesso de rotação. É criado a partir do evento 1014 sempre que o veículo não esteja em movimento.
- Evento 0141: Identifica rotações acima do valor máximo estipulado, quer o veículo esteja parado ou em movimento.
- Evento 1005: Regista as acelerações longitudinais bruscas, para velocidades inferiores a 10km/h.
- Evento 1007: Regista as desacelerações longitudinais bruscas, para velocidades superiores a 10km/h.
- Evento 1009: Regista situações de funcionamento do motor, superior a 3 minutos, desde que a viatura se encontre imobilizada.
- Evento 1010: Regista a activação do *kick-down*, acima das 1500RPM.
- Evento 1014: Regista as situações de rotação excessiva com a viatura imobilizada.
- Evento 1015: Regista as desacelerações longitudinais bruscas, para velocidades inferiores a 10 km/h.
- Evento 1026: Regista as vezes que o pedal de travão é accionado, desde que a viatura tenha velocidade superior a 1km/h.
- Evento 1058: Regista as situações de velocidade superior a 50km/h.
- Evento 1060: Regista as situações de velocidade superior a 60km/h.
- Evento 1062: Regista as situações de velocidade superior a 80km/h.
- Evento 1065: Regista as rotações inferiores a 500 RPM a uma velocidade superior a 25km/h.

- Evento 1067³: Regista as rotações acima do valor máximo estipulado, sempre que o veículo esteja a uma velocidade superior a 1km/h.
- Evento 1069: Regista o funcionamento do motor, superior a 3 minutos, com a viatura imobilizada e o ar condicionado ligado.
- Evento 1070: Regista as acelerações longitudinais bruscas, para uma velocidade superior a 10km/h.

Naturalmente, nos modelos de Carreiras e Viaturas, este valor resultará da média relativa ao total de viagens que englobe os eventos de todos os motoristas que tenham efectuado o serviço numa determinada carreira ou em determinada viatura, respectivamente.

O interesse destas variáveis é particularmente elevado, visto que poderão indicar quais os eventos de condução que prejudicam, ou não, o consumo de combustível verificado. Através destes resultados seria possível actuar junto dos motoristas no sentido de minorar os eventos que levassem de forma mais concreta ao aumento de consumo, através das acções de monitorização formativa já em vigor na empresa.

4.4.5. A política da empresa de transporte

4.4.5.1 Número médio de monitorizações formativas de motoristas

O número médio de monitorizações formativas de motorista foi calculado através da média do número de monitorizações dos vários motoristas que circularam numa determinada carreira ou num determinado veículo. A expressão que permite calcular esta variável, relativamente a uma determinada carreira ou viatura, é então a seguinte:

$$\text{NúmeroMonitorizações}_{\text{médio}} = \frac{\sum \text{NúmeroMonitorizações}}{\text{NúmeroTotalViagens}} \quad (16)$$

Novamente, esta variável só é assim calculada para os modelos de Carreiras e Viaturas visto que no modelo de Motoristas a variável toma o valor correspondente a cada motorista.

A introdução desta variável pretende servir como indicador do resultado que as acções de monitorização formativa com os motoristas têm no consumo de combustível. Assim, espera-se que para motoristas que tenham mais presenças em acções deste tipo apresentem uma melhor

³ Este evento é registado para rotações superiores a 1900rpm em veículos articulados e superiores a 2000rpm nos restantes.

eficiência ao nível do consumo de combustível e uma menor percentagem de ocorrência nos eventos de condução controlados.

4.5 Limitações na recolha de dados

Uma vez que não foi possível utilizar todos dados obtidos é necessário referir-se quais as limitações dos dados de que se dispõe para criação dos modelos.

4.5.1 Consumo de combustível médio

A forma como é obtido o valor do consumo de combustível diário na RL baseia-se na verificação da quantidade de combustível que, em cada dia, é necessário repor para encher totalmente o depósito. Assim, é possível saber a quantidade de combustível consumido, num determinado dia, por cada veículo. No entanto, este processo implica várias fontes de erro na medição do consumo, o que poderá ser bastante prejudicial à construção de modelos de regressão, uma vez que o consumo médio será utilizado como a variável dependente dos modelos. Este procedimento de determinação do consumo diário é tanto menos fiável quanto menor o tamanho do depósito de combustível de cada veículo, já que, para pequenas diferenças no que se refere ao enchimento do depósito na totalidade (processo com uma precisão reduzida), o erro relativamente ao consumo diário será maior. Além do erro associado ao depósito de combustível, há também que ter em conta o erro associado ao registo dos valores do conta-quilómetros do veículo, que, sendo inseridos manualmente, levam, muitas vezes, à obtenção de valores de consumo muito improváveis ou até mesmo impossíveis, tal como foi detectado na amostra de consumos recolhida. Por esta razão, foram criados filtros em que se eliminaram dos dados recolhidos valores que se consideraram fora da realidade, para que não entrassem no cálculo do valor de consumo de combustível médio obtido. Estes filtros foram definidos em colaboração com a Rodoviária de Lisboa, de acordo com o que a própria empresa efectua nas suas análises internas de consumo de combustível.

Os limites de cada filtro aplicados, de acordo com a tipologia de veículo, são os seguintes:

- Tipologia Mini: Eliminados dados fora do intervalo de 12 a 30 L/100km;
- Tipologia Mini: Eliminados dados fora do intervalo de 20 a 50 L/100km;
- Tipologia Mini: Eliminados dados fora do intervalo de 30 a 65 L/100km;
- Tipologia Mini: Eliminados dados fora do intervalo de 45 a 90 L/100km.

4.5.2 Variáveis de inclinação

No desenvolvimento de variáveis relativas à inclinação de cada carreira, optou-se por criar variáveis que considerassem intervalos de inclinação, em módulo (exemplificando, criou-se uma variável para declives superiores a 5% e outra com declives inferiores a 5%) independentemente do sentido do movimento (ascendente ou descendente). Decidiu-se unir os dados com inclinações em ambos os sentidos, uma vez que a mesma carreira tem de percorrer ambos os sentidos do percurso. Considerou-se este procedimento como aquele que permite, de forma mais simples, fazer uma avaliação do peso destas variáveis a partir de um modelo de regressão linear, uma vez que a distinção entre sentidos de deslocamento iria obrigar a criar mais variáveis que reflectissem o sentido do movimento em cada carreira, o que, na análise a um grande número de carreiras, não se considerou adequado.

4.5.3 Número de passageiros transportados

A informação relativa ao número de passageiros a ser transportados numa viatura ainda não é recolhida nas viaturas da RL. Desta forma, não se pôde criar esta variável, apesar de se considerar que poderia ter bastante interesse, até mesmo para outros tipos de análise que se pretendesse desenvolver como, por exemplo, ao nível da eficiência económica das viagens ou mesmo numa melhor gestão das paragens das várias carreiras.

Em princípio, num futuro próximo, com a introdução nos veículos de um novo sistema de bilhética sem contacto, à semelhança do que já ocorre, por exemplo, na operadora de transporte *Carris*, será possível obter este tipo de dados. Ainda assim, saber exactamente o número de passageiros transportados, em cada momento, na viatura, afigura-se bastante complicado, já que não existem previsões para a introdução de algum tipo de equipamento que identifique a saída dos passageiros da viatura.

4.5.4 Amostra no modelo de veículos

Em relação ao modelo de Viaturas, uma vez que foram tomadas como indispensáveis as variáveis relativas aos eventos de condução, o número da amostra para o estudo que se realizou apresenta uma dimensão muito inferior ao da totalidade de veículos em serviço por parte da RL. Isto deve-se ao facto de apenas 1/3 das viaturas estarem equipadas com o equipamento FM200, pelo que se tornou impossível obter valores relativos aos eventos de condução para muitas das viaturas. Uma vez que estas variáveis foram consideradas como muito interessantes na análise que se pretendia realizar e o número de viaturas equipadas era, ainda assim, suficiente à obtenção dos modelos através do método de regressão linear, optou-se por executar os modelos desta forma. Por outro

lado, esta opção levou a que no modelo de Viaturas acabasse por não existir nenhum caso de viaturas Midi, o que provavelmente inviabilizará o uso deste modelo para obter resultados para viaturas desta tipologia.

5. Apresentação e descrição de resultados

Os modelos desenvolvidos foram obtidos a partir do software SPSS utilizando o método de regressão *stepwise* e vão ser, em seguida, apresentados e interpretados, para cada um dos modelos. O método *stepwise* foi o seleccionado uma vez que foi aquele com que, sistematicamente, se obtiveram resultados com níveis de ajustamento muito próximos dos maiores alcançados, e ainda em que a selecção de variáveis aparentou ser mais criteriosa, uma vez que através dos outros dois métodos eram seleccionadas bastantes variáveis o que originava problemas de colinearidade. Os resultados aqui apresentados são aqueles que se consideraram como sendo os que melhor caracterizam o consumo de combustível após várias experimentações em cada modelo e também depois de feita a limpeza de dados recolhidos, pelo que se apresentam os modelos finais.

5.1 Resultados do modelo de carreiras

5.1.1 Sumário do modelo obtido

A expressão algébrica obtida para o modelo de carreiras foi a seguinte:

$$\begin{aligned} \log(\text{Consumption}) = & 1,741 - 0,012\text{NMonit} - 0,005\text{ComSpeed} - \\ & -0,004\text{VeicMini} + 1,173 \times 10^{-5} (\text{ISup}5\%)^2 + 0,019 \log(\text{MaxDistStops}) + \\ & + 0,48\text{Ev}0140 + 0,002\text{VeicArtic} - 0,001\text{VeicMini} \end{aligned} \quad (17)$$

Em seguida apresenta-se, na Tabela 6, o quadro síntese do modelo desenvolvido para a amostra de carreiras recolhida.

Tabela 6 - Sumário do modelo de Carreiras

R	R ²	R ² ajustado	Erro padrão da estimativa
,986	,972	,970	,01554

O valor de R² obtido para este modelo de 0.972 revelou-se muito satisfatório (pelo que explica 97,2% da variância dos dados que geraram o modelo).

O modelo apresenta 8 variáveis independentes explicativas do consumo esperado de uma determinada carreira, que se apresentam em seguida:

1. VeicMini - Percentagem de percursos feitos com veículo de tipologia Mini;
2. VeicMidi - Percentagem de percursos feitos com veículo de tipologia Midi;
3. (ISup5%)² - Quadrado da percentagem de percursos com inclinação superior a 5%;
4. VeicArtic - Percentagem de percursos feitas com veículo de tipologia Articulado;
5. Ev1040 - Percentagem de ocorrência do evento 1040 (carregamento do ar em excesso de rotação);
6. Log(MaxDistStops) - Logaritmo da distância máxima entre paragens;
7. NMonit - Número médio de monitorizações de motoristas;
8. ComSpeed - Velocidade comercial média.

Tendo-se utilizado 8 variáveis para descrever a variável dependente, a dimensão da amostra satisfaz, com uma boa margem, as condições mínimas referidas anteriormente, que sugeriam 5 observações por cada variável independente (TABACHNICK E FIDELL, 2007) em que a dimensão mínima da amostra seria de 40 observações e a presente amostra tem 87 observações. Ainda assim, a amostra recolhida situa-se um pouco abaixo dos valores mínimos sugeridos como indicados (em que para 8 variáveis independentes é sugerida a utilização de uma amostra com 114 observações).

Na tabela da análise ANOVA, Tabela A5 dos anexos, obteve-se um valor de $F=342,747$ com 8 e 78 g.l.. Esta estatística de teste tem associada um $p\text{-value} = 0.000$ pelo que se pode rejeitar H_0 em favor de H_1 , portanto o modelo é altamente significativo.

Na Tabela 7, apresenta-se a caracterização dos coeficientes de regressão atribuídos a cada uma das variáveis independentes.

Tabela 7 – Coeficientes estatísticos das variáveis do modelo de Carreiras

Modelo	Coeficientes não standardizados		Coeficientes standardizados	t	Sig.	Estatísticas de colinearidade	
	B	Erro Padrão	Beta			Tolerância	VIF
Constante	1,741	,024		71,494	,000		
Número de monitorizações	-,012	,003	-,117	-4,201	,000	,458	2,183
Velocidade Comercial	-,005	,001	-,263	-7,788	,000	,311	3,220
Tipologia Mini	-,004	,000	-,888	-42,120	,000	,798	1,253
Quadrado da percentagem de inclinação superior a 5%	1,173E-5	,000	,104	3,713	,000	,450	2,224
Logaritmo da distância máxima entre paragens	,019	,008	,070	2,269	,026	,368	2,719
Percentagem de ocorrência do evento 0140	,048	,021	,049	2,308	,024	,795	1,257
Tipologia Articulada	,002	,000	,238	12,053	,000	,911	1,098
Tipologia Midi	-,001	,000	-,167	-8,565	,000	,933	1,071

Da sua observação, é possível constatar que cada variável afecta significativamente a variável dependente para um nível de significância de $\alpha=0.05$, pelo que se optou por manter estas 8 variáveis independentes no modelo final.

5.1.2 Diagnósticos de multicolinearidade

Também através da análise da Tabela 7, é possível verificar que nenhuma das variáveis seleccionadas apresenta valores de VIF superiores a 5, o que permite concluir que não foram detectados fenómenos de multicolinearidade entre as variáveis independentes.

5.1.3 Análise de resíduos

A análise de resíduos é fundamental para a validação dos pressupostos de regressão linear pelo que se apresentam, em anexo, tabelas estatísticas e gráficos que permitem tirar as conclusões sobre a validade do modelo desenvolvido.

Através da observação da Tabela A7, em anexo, é possível concluir que os resíduos apresentam uma média igual a zero e ainda que os seus extremos, negativo e positivo, se distribuem de forma aproximadamente semelhante em torno da média.

Através da Figura A1, em anexo, uma vez que a maioria dos pontos se encontram aproximadamente em cima da diagonal principal, pode-se concluir que, os resíduos apresentam distribuição normal.

O histograma de resíduos, Figura A2, em anexo, apresentado permite, também, verificar a aproximação dos resíduos a uma distribuição normal.

Como se pode observar a partir da Figura A3 e A4, em anexo, não existem tendências não lineares assinaláveis e os resíduos distribuem-se aleatoriamente em torno de 0, sem nenhuma tendência detectável. Como conclusão destas análises, podem-se classificar os resíduos como homocedásticos.

5.1.4 Testes de normalidade

Outra forma de verificar a normalidade dos resíduos é através de testes paramétricos. As tabelas A8 e A9, em anexo, apresentam os resultados dos testes de normalidade de resíduos de Kolmogorov-Smirnov e Shapiro-Wilk, respectivamente.

Para o teste de Kolmogorov-Smirnov, na Tabela A8, sendo o $p\text{-value}=0.977$ não se rejeita a hipótese de que a variável *res_1* segue distribuição normal para os habituais níveis de significância ($\alpha=0.01$, $\alpha=0.05$ e $\alpha=0.10$).

De acordo com o teste Shapiro-Wilk, na Tabela A9, pode afirmar-se que os resíduos apresentam distribuição normal ($p=0.502>0.050$).

Podem detectar-se outliers através de um teste exacto uma vez que cada um dos resíduos *studentized deleted* possui uma distribuição t-student com $(n-p-1)$ g.l. Assim para testar a hipótese de que cada resíduo “Não é um outlier” (H_0) vs. “É um outlier” (H_1) pode-se calcular o p-value associado a cada um dos valores da variável *sdr_1*. Para um nível de significância $\alpha=0.05$, rejeita-se H_0 de que a observação j não é um outlier se $p\text{-value}\leq 0.05$. Desta forma foi possível identificar como outliers as carreiras: 305, 322, 342 e 361.

Para se analisar a influência de cada uma das observações sobre os valores estimados pode-se recorrer a um gráfico de dispersão entre os DfFits e os valores estimados. Através da Figura A5 em anexo, é possível concluir que nenhuns dos anteriores outliers detectados apresentam um p-value

muito próximo do limite de significância, e tendo em conta que a remoção de um outlier deve ser feita de forma muito cuidada, para não excluir eventuais fenómenos que este transmita ao modelo, optou-se por não remover estes outliers detectados.

5.2 Resultados do modelo de motoristas

5.2.1 Sumário do modelo obtido

A expressão algébrica obtida para o modelo de motoristas foi a seguinte:

$$\begin{aligned} \log(\text{Consumption}) = & 1,965 - 0,145 \log(\text{ComSpeed}) - 0,005 \text{VeicMini} - \\ & -0,002 \text{VeicStand} - 0,002 \text{VeicMidi} + 0,002 I \text{ sup } 5\% - \\ & -0,002 \text{AvrLenght} + 0,003 \text{Ev1007} \end{aligned} \quad (18)$$

Em seguida apresentam-se os quadros síntese do modelo desenvolvido para a amostra de motoristas recolhida:

Tabela 8 – Sumário do modelo de Motoristas

R	R ²	R ² ajustado	Erro padrão da estimativa
,958 ^a	,917	,916	,01750

A Tabela 8 demonstra que o modelo obtido apresenta um valor de R² de 0.917, o que representa 91.7% de explicação da variância total da variável dependente.

Através da Tabela A11, em anexo, apresenta-se a análise ANOVA através da qual é possível identificar que o valor de F é 758,309 com 7 e 480 g.l.. Esta estatística de teste tem associada um p-value = 0.000 pelo que podemos rejeitar H0 em favor de H1, portanto o modelo é altamente significativo.

O modelo é composto por 7 variáveis independentes que, em seguida, se enumeram:

1. Ev1007 – Percentagem de ocorrência do evento 1007 (desacelerações longitudinais bruscas);
2. VeicStand – Percentagem de condução de viaturas de tipologia Standard;
3. VeicMini – Percentagem de condução de viaturas de tipologia Mini;
4. VeicMidi – Percentagem de condução de viaturas de tipologia Midi;
5. ISup5% – Percentagem percursos com inclinação superior a 5%;

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA**

6. log(ComSpedd) – Logaritmo da velocidade comercial;
7. AvrLenght – Comprimento médio de percurso.

Neste modelo utilizaram-se, 7 variáveis para descrever a variável dependente, pelo que o número da amostra satisfaz, largamente, as condições mínimas referidas anteriormente que sugeriam 5 observações por cada variável independente (a dimensão mínima da amostra seria de 35 observações e a presente amostra tem 488 observações). Desta forma, também se verifica o valor mínimo aconselhado de variáveis sugerido em TABACHNICK E FIDELL, 2007.

Através da observação da Tabela 9, confirma-se que todas as variáveis afectam significativamente a variável dependente para um nível de significância de $\alpha=0.05$.

Tabela 9 – Coeficientes estatísticos das variáveis do modelo de Motoristas

Modelo	Coeficientes não standardizados		Coeficientes standardizados beta	t	Sig.	Estatísticas de colinearidade	
	B	Erro padrão	Beta			Tolerância	VIF
Constante	1,965	,024		82,230	,000		
Logaritmo da velocidade comercial	-,145	,020	-,161	-7,360	,000	,360	2,779
Percentagem de condução de tipologia Mini	-,005	,000	-,845	-55,150	,000	,736	1,359
Percentagem de condução de tipologia Standard	-,002	,000	-,561	-33,719	,000	,624	1,604
Percentagem de condução de tipologia Midi	-,002	,000	-,325	-21,222	,000	,735	1,360
Percentagem percursos com inclinação superior a 5%	,002	,000	,310	14,390	,000	,373	2,685
Comprimento médio de percurso por motorista	-,002	,000	-,132	-4,906	,000	,237	4,219
Percentagem de ocorrência do evento 1007	,003	,001	,077	5,689	,000	,943	1,060

5.2.2 Diagnósticos de multicolinearidade

Consultando novamente a Tabela 9, é possível indicar que nenhuma das variáveis seleccionadas apresenta valores de VIF superiores a 5, pelo que se pode concluir que não foram detectados fenómenos de multicolinearidade entre as variáveis independentes.

5.2.3 Análise de resíduos

Através da análise da tabela de estatísticas de resíduos e um conjunto de gráficos de caracterização de resíduos. Fazendo a interpretação rápida da Tabela A13, em anexo, de estatísticas de resíduos, constata-se que a média dos resíduos é zero e que estes se distribuem de forma idêntica em torno da média.

Sendo este o modelo com a maior amostra, é neste que se destaca uma definição mais perfeita da diagonal principal, através da Figura A6, em anexo, pelo que se pode concluir que os resíduos apresentam distribuição normal. Através da observação do histograma de resíduos, Figura A7 dos anexos, é perceptível uma aproximação dos resíduos a uma distribuição normal.

Os gráficos de dispersão de resíduos, figuras A8 e A9 dos anexos, permitem verificar alguns pressupostos de regressão linear necessários à sua validação. Como se pode observar a partir destas figuras, não existem tendências não lineares assinaláveis e os resíduos distribuem-se aleatoriamente em torno de 0, sem nenhuma tendência clara. Em ambas as figuras é possível constatar que os resíduos se distribuem de forma mais ou menos aleatória em torno de zero e, além disso, não é detectável a presença de outliers. Com estas análises, fica assegurada a homocedasticidade dos resíduos neste modelo.

5.2.4 Testes de normalidade

Mais uma vez, para se proceder a testes de normalidade utilizou-se os testes de Kolmogorov-Smirnov e Shapiro-Wilk.

O teste de Kolmogorov-Smirnov, Tabela A14, em anexo, permite averiguar se os erros seguem ou não distribuição normal. Sendo o p-value=0.784 não se rejeita a hipótese de que a variável res_1 segue distribuição normal para os habituais níveis de significância ($\alpha=0.01$, $\alpha=0.05$ e $\alpha=0.10$).

De acordo com o teste Shapiro-Wilk, Tabela A15, não se rejeita a hipótese de que os resíduos apresentem distribuição normal ($p=0.100 > 0.050$).

Em seguida pretende-se detectar outliers através do teste exacto já utilizado anteriormente. Para testar a hipótese de que cada resíduo “Não é um outlier” (H0) vs. “É um outlier” (H1) calcula-se o p-

-value associado a cada um dos valores da variável *sdr_1*. Para um nível de significância $\alpha=0.05$, rejeita-se H_0 de que a observação *j* não é um outlier se $p\text{-value} \leq 0.05$. Desta forma foi possível identificar como outliers os motoristas: 67603, 78782, 81935, 90639, 90647, 90738, 91181, 91876, 92445, 92825, 96875 e 96909.

Para se analisar a influência de cada uma das observações sobre os valores estimados pode-se recorrer a um gráfico de dispersão entre os DfFits e os valores estimados, ver Figura A10, em anexo.

Uma vez que os outliers detectados não demonstram ser muito influentes para a regressão e o seu *p-value* se situa próximo do limite de significância, optou-se por não remover estes outliers do modelo.

5.3 Resultados do modelo de viaturas

5.3.1 Sumário do modelo obtido

A expressão algébrica obtida para o modelo de viaturas foi a seguinte:

$$\log(\textit{Consumption}) = -1,737 + 0,776 \log(\textit{VeicMass}) - 0,006 \textit{ComSpeed} - 0,003 \textit{AvrgLenght} + 0,017 \textit{Ev1067} + 0,195 \log(\textit{DriversAge}) \quad (19)$$

Em seguida apresenta-se, na Tabelas 10, o quadro síntese do modelo desenvolvido para a amostra de viaturas recolhida:

Tabela 10 - Sumário do modelo de Viaturas

R	R ²	R ² Ajustado	Erro padrão da estimativa
,977 ^a	,954	,952	,02393

O valor de R^2 obtido para este modelo revelou-se muito significativo sendo de 0.954, explicando 95,4% da variância total da variável dependente.

Fazendo a interpretação do quadro da ANOVA (Tabela A17 dos anexos) obteve-se um valor de $F=413,370$ com 5 e 99 g.l. Esta estatística de teste tem associada um *p-value* = 0.000 pelo que podemos rejeitar H_0 em favor de H_1 , portanto o modelo é altamente significativo.

O modelo apresenta 5 variáveis independentes explicativas do consumo esperado de uma determinada carreira:

1. $\log(\textit{DriversAge})$ – Logaritmo da idade média de motoristas a circular;

2. Ev1067 – Percentagem de ocorrência do evento 1067 (rotações acima do valor máximo estipulado);
3. ComSpeed – Velocidade comercial;
4. log(VeicMass) – Logaritmo da massa;
5. AvrgLenght – Comprimento médio de percurso.

No modelo de viaturas foram utilizadas 5 variáveis para descrever a variável dependente, pelo que o número da amostra satisfaz as condições mínimas de dimensão da amostra de 35 observações. Além disso, tendo este modelo uma amostra de 105 observações considera-se que se cumpre o valor mínimo aconselhado segundo TABACHNICK E FIDELL, 2007, de que para 7 variáveis independentes se deveria ter uma amostra de 106 observações.

A Tabela 11 refere-se à caracterização dos coeficientes de regressão atribuídos a cada uma das variáveis independentes.

Tabela 11 – Coeficientes estatísticos das variáveis do modelo de Viaturas

Modelo	Coeficientes não standardizados		Coeficientes standardizados	t	Sig.	Estatísticas de colinearidade	
	B	Erro padrão	Beta			Tolerância	VIF
Constante	-1,737	,141		-12,311	,000		
Logaritmo da massa de cada viatura	,776	,020	,912	38,079	,000	,805	1,242
Velocidade comercial	-,006	,001	-,195	-4,735	,000	,273	3,666
Comprimento médio de percurso	-,003	,001	-,140	-3,325	,001	,261	3,831
Percentagem de ocorrência do evento 1067	,017	,004	,108	4,838	,000	,931	1,075
Logaritmo da idade média de motoristas	,195	,077	,056	2,516	,013	,927	1,079

Através da observação da Tabela 11, é possível verificar que cada variável afecta significativamente a variável dependente para um nível de significância de $\alpha=0.05$, pelo que se mantiveram as 5 variáveis seleccionadas.

5.3.2 Diagnósticos de multicolinearidade

Nenhuma das variáveis seleccionadas apresenta valores de VIF superiores a 5 (ver Tabela 11), pelo que se pode concluir que não foram detectados fenómenos de multicolinearidade entre as variáveis independentes.

5.3.3 Análise de resíduos

Analisando os valores de resíduos apresentados na Tabela A19, em anexo, conclui-se que os resíduos apresentam uma média igual a zero com extremos distribuídos de forma semelhante em torno dessa média.

Pela observação da Figura A11, em anexo, uma vez que a maioria dos pontos está sobre a diagonal principal, pode-se concluir que os resíduos apresentam, pelo menos aproximadamente, uma distribuição normal. Através da observação do histograma de resíduos, Figura A12, em anexo, é perceptível uma aproximação dos resíduos a uma distribuição normal.

Os gráficos de dispersão de resíduos, figuras A13 e A14, em anexo, permitem avaliar alguns pressupostos de regressão linear necessários à sua validação. Como se pode observar a partir destas figuras, não existem tendências não lineares assinaláveis e os resíduos distribuem-se aleatoriamente em torno de 0. Nestes dois gráficos é possível constatar que os resíduos se distribuem de forma mais ou menos aleatória em torno de zero e, além disso, não é detectável a presença de outliers. Com estas observações gráficas pode-se concluir que os resíduos são homocedásticos.

5.3.4 Testes de normalidade

Para se averiguar a normalidade dos resíduos, vai-se proceder aos testes de Kolmogorov-Smirnov e Shapiro-Wilk.

O teste de Kolmogorov-Smirnov, Tabela A20, em anexo, permite averiguar se os erros seguem ou não distribuição normal. Sendo o p-value=0.731 não se rejeita a hipótese de que a variável res_1 segue distribuição normal para os habituais níveis de significância ($\alpha=0.01$, $\alpha=0.05$ e $\alpha=0.10$).

De acordo com o teste Shapiro-Wilk, Tabela A21, em anexo, conclui-se que os resíduos apresentam distribuição normal ($p=0.574>0.050$). Utilizando o mesmo processo descrito anteriormente vai-se testar a hipótese de que cada resíduo “Não é um outlier” (H_0) vs. “É um outlier” (H_1). Desta forma foi possível identificar como outliers as viaturas: 705, 708, 743 e 798.

Para se analisar a influência de cada uma das observações sobre os valores estimados pode-se fazer um gráfico de dispersão entre os DfFits e os valores estimados (ver Figura A15). Uma vez que os anteriores outliers detectados não demonstram ser muito influentes para a regressão e ainda o facto de o seu p-value se situar próximo do limite de significância, optou-se por não remover estes outliers do modelo.

6. Discussão e interpretação de resultados

No presente capítulo, pretende-se desenvolver uma análise dos resultados que se obtiveram de forma conseguir-se extrair dos mesmos informações válidas relativamente ao tema que se pretende estudar, nomeadamente, a maneira como as variáveis seleccionadas conseguem explicar a variável dependente (o consumo de combustível em litros por 100km) e que ganhos se prevêem obter através de futuras alterações nas variáveis escolhidas para cada modelo.

Para isso, em seguida, irá proceder-se a uma análise relativa a cada modelo e também a cada variável independente seleccionada por modelo, ilustrando através de exemplos como é esperado, por cada modelo, que os consumos sejam alterados de acordo com a variação das variáveis independentes. Além disto, também se irá procurar relações entre variáveis desenvolvidas que se considere terem interesse no estudo em questão.

Também será feita a validação de resultados com base numa nova amostra, relativa aos meses de Junho de 2010 e Setembro de 2010. Com isto, serão desenvolvidos gráficos de dispersão relativos aos valores de consumo para cada caso das novas amostras recolhidas, comparando-os com os valores previstos pelos modelos desenvolvidos de acordo com as novas observações. Além de se apresentar o gráfico de relação entre os consumos previstos pelos modelos e os consumos observados, será também determinado o valor do ajustamento (R^2) do modelo em relação aos novos dados, verificando se estão próximos dos mesmos valores registados em relação à amostra a partir da qual cada modelo foi desenvolvido.

6.1 Modelo Carreiras

6.1.1 Considerações gerais

O modelo de carreiras foi aquele que, de entre os 3 modelos desenvolvidos, apresentou um maior valor de R^2 , conseguindo explicar 97,0% da variância total da variável dependente. Além disso, através dos dados utilizados para validação do modelo de regressão, conseguiu-se uma explicação da variância na ordem dos 95,6%, que se considera um valor bastante elevado.

O modelo de regressão obtido para o conjunto de carreiras em estudo tem como variável dependente o logaritmo do consumo, uma vez que se verificou que esta transformação permitia uma melhor correlação com as variáveis independentes que o explicavam.

6.1.2 Análise de variáveis

6.1.2.1 Tipologia de veículos

Para o modelo de carreiras, as variáveis seleccionadas que representam a tipologia de veículo foram a Tipologia Mini(%), Tipologia Midi(%) e Tipologia Articulado(%).

De acordo com as massas de cada tipologia, é possível constatar que a percentagem de tipologia Mini é a que leva a uma maior diminuição do consumo, seguida da percentagem de tipologia Midi. Por outro lado, o modelo desenvolvido reconhece que um aumento da percentagem de utilização de veículos articulados levará a um aumento nos valores de consumo expectáveis.

Através do gráfico de elasticidades (ver Figura A 15 em, anexo) é possível verificar que para qualquer das tipologias de veículo o consumo é inelástico, visto que a sua elasticidade apresenta valores inferiores a 1. Ainda assim, analisando os gráficos, é concluído que o logaritmo de consumo de combustível apresenta maiores elasticidades para as tipologias Mini e Articulado. No entanto, o coeficiente destas variáveis apresenta sinal oposto e, portanto, o aumento da tipologia Mini levará a uma diminuição do consumo (tal como o aumento da tipologia Midi) enquanto o aumento da tipologia Articulado conduz ao aumento do consumo de combustível.

6.1.2.2 Velocidade comercial

A velocidade comercial média é a variável independente que apresenta o segundo maior valor (absoluto) de coeficiente estandardizado. Neste caso, apresenta sinal negativo o que implica que o aumento desta variável independente se traduz numa espectável redução de consumo, como sugerido por UITP, 2009. Este resultado permite concluir que a velocidade comercial, além de ser um parâmetro de qualidade do serviço prestado ao cliente, é também fundamental no que refere aos consumos registados. Assim sendo, é fundamental verificar de que forma se pode actuar sobre a velocidade comercial, nomeadamente ao nível da definição de percursos de carreiras, número e distância entre paragens e a passagem por troços onde o tráfego seja muito intenso.

O gráfico de elasticidades (ver Figura A 16, em anexo) permite verificar que o consumo de combustível diminui com a velocidade comercial e que esta variável implica um dos mais fortes valores de elasticidade entre as variáveis seleccionadas pelo modelo, justificando a sua importância nas variações do valor de consumo expectável.

A título de exemplo, é possível concluir que para um aumento da velocidade comercial de 14km/h para 16km/h (que constitui uma variação de 14,3%), pode-se esperar uma diminuição de 2,2% no consumo de combustível da carreira em questão.

Através da análise do seguinte gráfico de dispersão, em que a cada ponto é atribuída a sua velocidade comercial e a distância média entre paragens, é fácil constatar que existe uma relação bastante considerável entre ambas as variáveis. Essa relação indica que a velocidade comercial aumenta consideravelmente com o aumento da distância média entre paragens, pelo que devem ser bem planeadas não só a localização das paragens mas também o seu número e distância a que se situam umas das outras, visto que se demonstrou que a velocidade comercial é um factor fundamental na previsão do consumo de combustível.

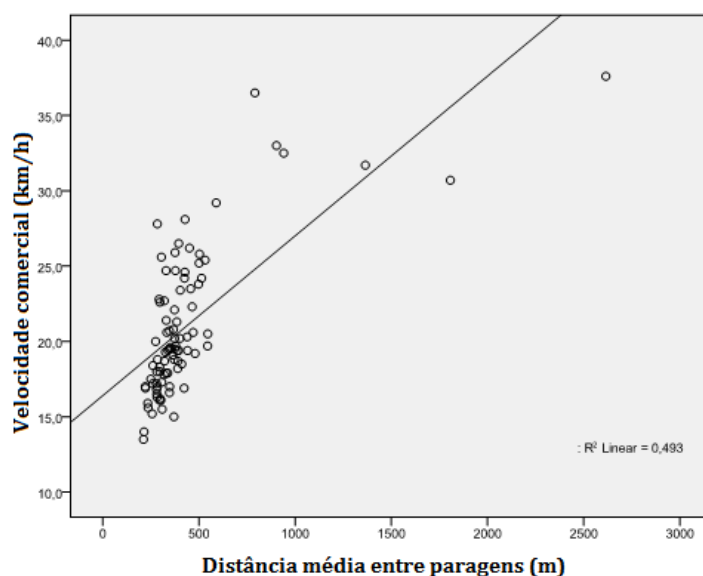


Figura 20 - Velocidade comercial vs Distância média entre paragens.

6.1.2.3 Número médio de monitorizações

A variável relativa ao número médio de monitorizações formativas de motoristas em cada carreira surge como a variável independente que apresenta o quarto maior coeficiente estandardizado (absoluto), apresentando um sinal negativo, o que significa que o seu aumento leva a uma redução do consumo de combustível.

Analisando o gráfico de elasticidade do consumo de combustível em relação ao número médio de monitorizações (ver Figura A 17, em anexo) conclui-se que a variável dependente apresenta uma elasticidade mais baixa do que relativamente às anteriores variáveis mas, ainda assim, é possível concluir que o aumento do número de monitorizações faz diminuir o consumo de combustível previsto.

Como exemplo, pode-se constatar que para um aumento do número de monitorizações formativas médio de 2,5 para 4 sessões (que constitui uma variação de 60%), pode-se esperar uma diminuição de 4,1% no consumo de combustível de uma carreira.

A presença desta variável neste modelo vem confirmar a importância que deve continuar a ser dada à formação em sessões de monitorização com os motoristas, para que se diminuam ao máximo as práticas erradas de condução e com isso se consigam obter melhores resultados quer ao nível do consumo de combustível como também da segurança e do conforto. Estes resultados suportam, portanto, o sucesso do esforço que tem sido desempenhado pela empresa neste sentido.

O gráfico seguinte ilustra a dispersão entre a variável número de monitorizações médio e a percentagem de ocorrência do evento 1067 - rotações acima do valor máximo estipulado. Através deste gráfico pretende-se verificar a influência das monitorizações na acumulação de eventos deste tipo.

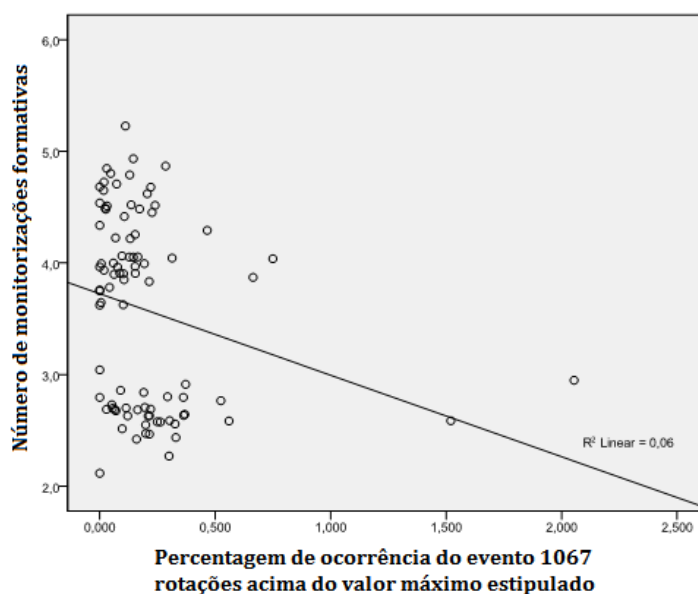


Figura 21 - Percentagem de ocorrência do evento 1067 vs Número médio de monitorizações formativas de motoristas

Através da observação do gráfico anterior, é possível detectar uma tendência, ainda que bastante ligeira, de que o aumento de número de monitorizações induza a uma diminuição da ocorrência do evento 1067 (evento detectado como influente no aumento dos consumos). Assim, demonstra-se directamente a relação entre as sessões de monitorização formativa e os eventos de condução erróneos por parte dos motoristas, comprovando-se que estas iniciativas resultam no desenvolvimento do desempenho dos motoristas da empresa.

6.1.2.4 Percentagem de percurso com inclinação superior a 5%

A variável relativa aos percursos com uma inclinação superior a 5% tem um coeficiente estandardizado de 0.104. Segundo o resultado que se obteve, um aumento da percentagem de

percurso com inclinações superiores a 5% traduz-se num aumento do consumo registado. Este resultado já era esperado, uma vez que vários estudos também apontam nesse sentido (HENSHER, 2003). No entanto, para se alterar o valor desta variável em cada carreira, seria necessária uma revisão e alteração dos seus percursos, algo que à partida não fará sentido para as carreiras já existentes. Este resultado deve então servir para se ter em consideração que, quer na gestão das carreiras existentes quer na definição de novos percursos, a percentagem de inclinação superior a 5% é um factor relevante no que refere ao consumo esperado.

A Figura A 18, em anexo, representa a elasticidade do consumo de combustível face à percentagem de percurso com inclinação superior a 5%. Pela sua observação, verifica-se que o consumo não é muito sensível a variações da percentagem de inclinação superior a 5%. Ainda assim, o efeito desta característica topográfica não é desprezável no cálculo do consumo de combustível. Ainda assim, utilizando um exemplo mais concreto, é possível concluir que para uma variação da percentagem de inclinação de 25.1% para 30% se obteria um aumento do consumo de 0.7%.

Em relação a esta variável, considerou-se oportuno avaliar a sua relação com os 3 centros de actividade de transporte da RL. Para isso calculou-se a matriz de correlação entre o consumo de combustível e os 3 centros de actividade de transporte e a matriz de relação entre a inclinação superior a 5% e as variáveis relativas a cada CAT.

Tabela 12 - Correlação entre Consumo de combustível e quadrado da inclinação superior a 5% com os Centros de Actividade de Transporte

	(I>5%)²	Consumo
CAT5	0,74	0,31
CAT6	-0,46	-0,39
CAT7	-0,35	0,06

Da observação deste conjunto de tabelas, extraídas das respectivas matrizes de correlação, é possível identificar algumas semelhanças entre ambas, o que vem também reforçar a relação entre o consumo e a inclinação geral do percurso de uma carreira. Desde logo, é possível identificar que, em ambas as tabelas, as variáveis analisadas, consumo e inclinação superior a 5%, apresentam uma correlação positiva com o CAT 5. Isto traduz que no CAT5 as carreiras têm um percurso mais acidentado o que, por sua vez, acaba por influenciar a um aumento dos níveis de consumo.

Em sentido inverso, constata-se que o CAT6 apresenta uma correlação negativa quer com o consumo quer com a variável de inclinação superior a 5%, o que acaba por indicar, novamente, que,

sendo os percursos das carreiras do CAT6 menos acidentados, estão-lhes associadas viagens com consumos mais baixos.

Por último, o CAT 7 é aquele que apresenta valores menos representativos (mais baixos em valor absoluto) quer na tabela que diz respeito ao consumo quer na que se refere à percentagem de inclinações superiores a 5%. Este resultado permite concluir que o CAT7 apresentará valores médios destas duas variáveis face aos outros 2 centros de actividade de transporte.

É possível ainda confirmar estes mesmos resultados através dos valores médios de ambas as variáveis obtidos para cada centro de actividade, apresentados na próxima tabela.

Tabela 13 - Tabela de inclinação superior a 5% e consumo de combustível por Centro de Actividade de Transporte

	Inclinação Superior a 5% (%)	Consumo (L/100km)
CAT5	38.97	47.30
CAT6	18.72	40.41
CAT7	20.87	45.10

6.1.2.5 Distância máxima entre paragens

A distância máxima entre paragens é uma variável que apresenta um valor de coeficiente estandardizado de +0.070, tendo sido transformada para o seu logaritmo de forma a obter uma melhor correlação com a variável dependente. Este coeficiente indica que um aumento das distâncias máximas entre paragens levará também a um aumento do consumo esperado, tal como sugerido em SIMÕES, 2005.

Uma vez que a variável distância máxima entre paragens foi transformada no seu logaritmo, a elasticidade corresponde ao coeficiente de regressão obtido, que neste caso é 0.019. O valor da elasticidade é muito reduzido, pelo que o consumo de combustível não aparenta ser muito sensível a variações da distância máxima entre paragens.

Além de se relacionar esta variável com o consumo, considerou-se que também seria interessante verificar como se relacionaria com a velocidade comercial, variável que denota grande importância na previsão do consumo, tal como já se tinha efectuado entre a velocidade comercial e a distância média entre paragens. O próximo gráfico de dispersão que relaciona o logaritmo da distância máxima entre paragens e a velocidade comercial comprova que estas variáveis apresentam uma correlação positiva, podendo-se portanto admitir que o aumento da distância máxima entre paragens também tenderá para um aumento da velocidade comercial. Esta relação vem provar a

importância que deve ser dada ao traçado dos itinerários das carreiras e ao posicionamento das suas paragens.

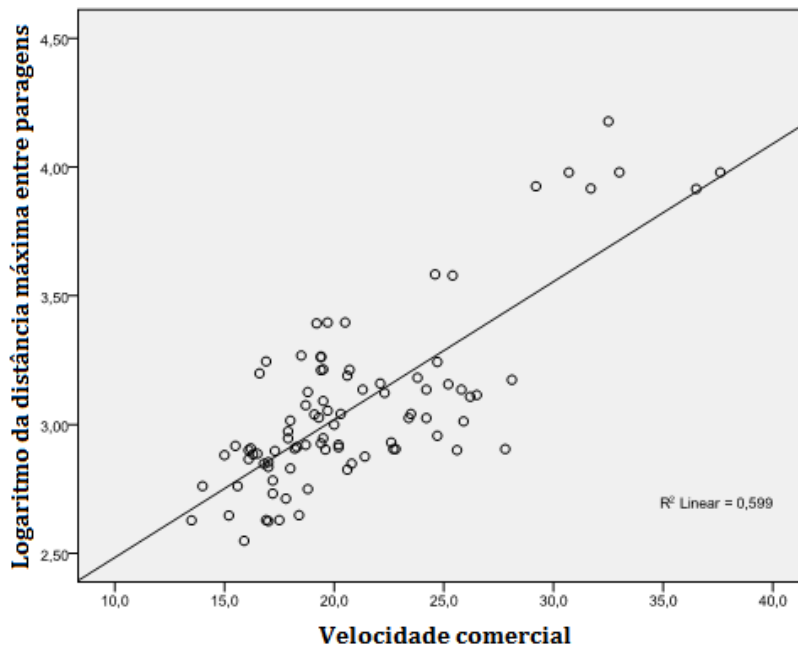


Figura 22 – Velocidade comercial (km/h) vs Logaritmo da distância máxima entre paragens.

6.1.2.5 Percentagem de ocorrência do evento 1040 (carregamento do ar em excesso de rotação)

A ocorrência de carregamento do ar em excesso de rotação foi detectada como tendo relevância no aumento do consumo esperado. Apesar de apresentar o mais baixo valor de coeficiente estandardizado (+0.049), esta variável ainda é tida, estatisticamente, como significativa para a determinação da variável dependente, o consumo de combustível. Este é um resultado com bastante interesse uma vez que, através do esforço feito pela empresa em registar vários tipos de eventos de condução, foi possível concluir que o carregamento do ar em excesso de rotação apresenta uma relação com o aumento do consumo de uma viatura e, portanto, deve-se intervir junto dos motoristas de forma a que estes estejam mais alerta para esta situação, para se obterem ganhos na eficiência energética dos serviços prestados.

Em anexo, apresenta-se o gráfico de elasticidade (ver Figura A 19) do consumo de combustível em relação à percentagem de ocorrência do evento 1040 - carregamento do ar em excesso de rotação. Pela observação do gráfico facilmente se conclui que o consumo é pouco sensível face a alterações da percentagem de ocorrência do evento 0140. Ainda assim, a título exemplificativo, estima-se que uma variação da percentagem de ocorrência do evento 0140 de 0,056% para 0,010% leva a uma diminuição de consumo de 0,5%.

6.1.2.6 Validação do modelo de Carreiras

Para a validação do modelo de carreiras conseguiu-se obter uma nova amostra com a dimensão de 71 carreiras. O valor de ajustamento da nova amostra ao modelo foi de $R^2=0.953$, sendo que o nível de ajustamento obtido para os dados que originaram o modelo foi de $R^2=0.970$. Desta forma, pode-se concluir que os resultados obtidos através do modelo desenvolvido para a nova amostra recolhida apresentam uma explicação da variância muito alta, validando assim a utilização do modelo para efeitos de estimação de consumo de combustível de uma carreira. Em seguida, através da Figura 23, facilmente se pode distinguir que, para valores de consumo de combustível médio/altos, tipicamente associados carreiras com percursos efectuadas por viaturas de tipologia Standard, a recta de valores previstos intersecta a recta de maior tendência de valores esperados, o que é explicável devido à maior parte das carreiras ser efectuada por viaturas dessa tipologia.

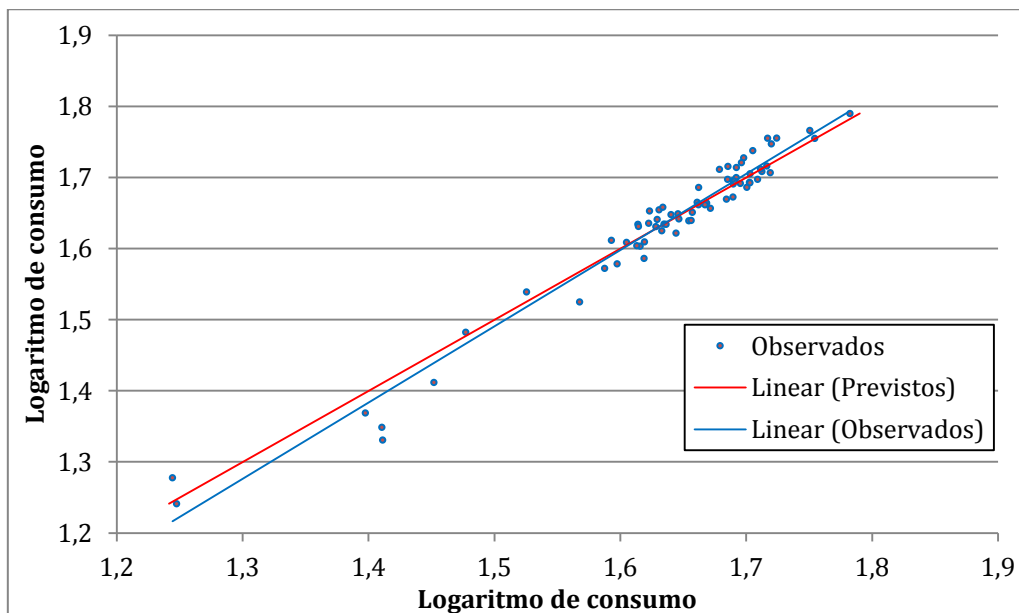


Figura 23 - Comparação entre valores previstos de logaritmo de consumo pelo modelo de Carreiras e valores observados da amostra de validação.

6.2 Modelo Motoristas

6.2.1 Considerações gerais

O modelo de motoristas apresenta uma explicação da variância total da variável dependente de 91,7%, sendo que pelos dados de validação do modelo, se conseguiu uma explicação da variância na ordem dos 84,6%. Estes valores, apesar de inferiores aos apresentados em ambos os outros modelos são igualmente bastante bons. Esta menor explicação obtida pelo modelo de motoristas poderá estar relacionada com o facto de a amostra obtida ser muito superior à dos outros dois

modelos (com cerca de 500 observações), mas também com o factor humano que está associado ao modelo em questão, o que, só por si, poderá implicar uma maior variabilidade de valores face aos outros modelos.

O modelo de regressão obtido para o grupo de motoristas analisado apresenta como variável dependente o logaritmo do consumo, uma vez que se verificou que esta transformação obtinha uma melhor correlação com as restantes variáveis independentes.

6.2.1 Análise de variáveis

6.2.1.1 Tipologia de veículos

As variáveis seleccionadas que representam a tipologia de veículo foram a **Tipologia Mini(%)**, **Tipologia Midi(%)** e **Tipologia Standard(%)**.

De acordo com as massas de cada tipologia, é possível constatar que a percentagem de tipologia Mini é a que apresenta o maior índice de coeficiente estandardizado (em valor absoluto) do consumo, seguida da percentagem de tipologia Standard, o que se revelou inesperado, uma vez que para viaturas de tipo Midi os consumos médios são inferiores ao das viaturas Standard e, portanto, poderia ser de prever que o coeficiente estandardizado relativo ao tipo Midi fosse superior. Uma das razões para este resultado menos de acordo com o esperado poderá ser devido ao facto de o número de viaturas Midi ser diminuto e, portanto, a sua influência ser de mais difícil interpretação e adaptação ao modelo.

Através do gráfico de elasticidades (ver Figura A 20, em anexo), é possível identificar algumas semelhanças relativamente ao mesmo gráfico do modelo de Carreiras. Desde logo, a elasticidade relativa à tipologia Mini é novamente a que apresenta maior elasticidade. Deve-se ainda referir que também os valores de elasticidade para as tipologias de viaturas deste modelo têm uma ordem de grandeza e valores bastante semelhantes aos apresentados no modelo de Carreiras.

6.2.1.2 Percentagem percursos com inclinação superior a 5%

Esta variável de caracterização topográfica surge novamente como relevante para a determinação da variável dependente, surgindo neste modelo ainda com mais peso (é a variável com maior coeficiente standardizado depois das variáveis de tipologia de veículo).

A Figura A 21, em anexo, apresenta a elasticidade do consumo de combustível em relação à variação da percentagem média de inclinação superior a 5% para um motorista. Os valores de elasticidade muito semelhantes aos obtidos para o modelo de Carreiras, pelo que os comentários

serão análogos. Ainda assim, para uma variação na percentagem de inclinação superior a 5% de 37% para 25%, o consumo médio por motorista deverá apresentar uma diminuição de 5.3%.

6.2.1.3 Velocidade comercial

A velocidade comercial é a única variável, além das variáveis relativas ao tipo de veículo, que está presente nos 3 modelos desenvolvidos. Tal como nos outros modelos, a velocidade comercial volta a ser uma das variáveis mais significativas para a estimação do consumo através de um modelo de regressão. Neste caso, apresenta um coeficiente estandardizado de -0.161, pelo que a tendência se mantém, ou seja, o aumento do valor desta variável traduz-se numa redução do consumo esperado.

Uma vez que no modelo de motoristas a velocidade comercial foi transformada no seu logaritmo, a elasticidade corresponde ao coeficiente de regressão obtido, que neste caso é 0.145. Este é um valor considerável, comparando-o com os restantes obtidos para o modelo de Motoristas, pelo que a importância da inclusão da velocidade comercial no modelo é assim confirmada. É possível ainda prever que, para uma variação entre 25,7km/h para 19km/h o consumo de combustível aumenta cerca de 4,5%.

Tendo em conta a importância que a variável independente velocidade comercial apresentou nos vários modelos desenvolvidos, importa analisar de que formas esta poderá variar. Assim, considerou-se importante procurar outras variáveis com as quais esta apresentasse uma boa relação. Uma das variáveis com que a velocidade comercial apresentou uma relação considerável foi a percentagem de inclinação inferior a 5%. Segundo o gráfico obtido, para um aumento da percentagem de percurso com inclinação inferior a 5% a velocidade comercial tem uma tendência, bastante clara, para também aumentar, o que, como já se provou, é benéfico para uma melhoria da eficiência energética.

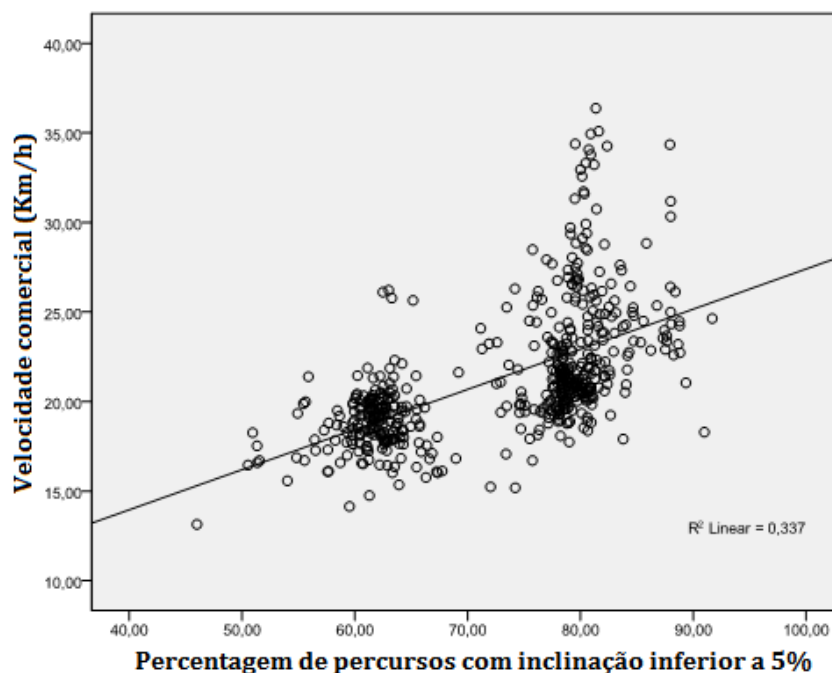


Figura 24 - Percentagem de percursos com inclinação superior a 5% vs Velocidade comercial.

6.2.1.4 Comprimento médio de percurso

O comprimento médio de percurso apresenta um coeficiente estandardizado de -0.132, o que significa que uma diminuição de percurso indicia um aumento do consumo de combustível. Este resultado poderá estar relacionado com o facto de as carreiras com percursos mais longos circularem a velocidades médias superiores através de vias com melhores condições e diferentes limites de velocidade, o que leva a uma melhor eficiência no consumo dispendido na deslocação.

A Figura A 22, em anexo, permite identificar a elasticidade que o comprimento médio de percurso induz no consumo médio de cada motorista, apresentando como máximo valor de elasticidade cerca de 0.045, voltando a ser das variáveis que menor elasticidade provocam ao consumo de combustível.

Esta variável foi introduzida no modelo com o intuito de verificar se o comprimento de percurso teria, ou não, alguma influência no registo de eventos de condução errados. Na grande maioria dos casos esta relação revelou-se inexistente, sendo que, nos gráficos de dispersão obtidos não era identificável qualquer tendência. Ainda assim, ao testar-se a influência do comprimento médio de percurso sobre o excesso de velocidade, foi possível constatar que entre ambos existe uma correlação positiva, o que sugere que os motoristas tendem a exceder mais vezes a velocidade (acima dos limites estabelecidos) quando os percursos são maiores. Esta conclusão poderá estar associada a diversos factores como a vontade do próprio motorista em viajar a uma maior velocidade (já foi identificado pela RL que alguns motoristas cumprem a última viagem diária a uma

velocidade superior àquela a que normalmente fazem o mesmo percurso noutras situações). Outra possível razão para esta efeito poderá advir da pressão exercida sobre o motorista por parte dos passageiros das viaturas, pressão essa que será, certamente, maior em viagens mais longas. Além disso, é natural que as carreiras mais longas sejam muitas das vezes aquelas que circulam por vias com melhores condições e com limites de velocidade maiores do que os verificados dentro de localidades, o que poderá explicar a relação deste evento com o comprimento médio.

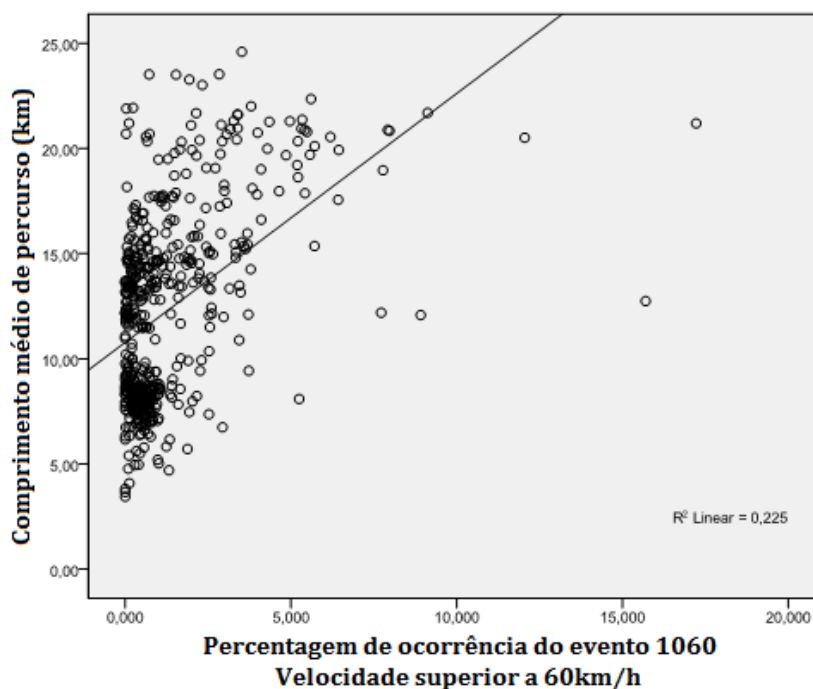


Figura 25 - Percentagem de ocorrência do evento 1060 vs Comprimento médio de percurso.

6.2.1.5 Percentagem de ocorrência do evento 1007 (desacelerações longitudinais bruscas)

O evento relativo à percentagem de ocorrência de desacelerações longitudinais bruscas (travagens) apresenta uma correlação positiva com o consumo, sendo que o seu coeficiente estandardizado é +0.077. Desta forma, este resultado sugere que o aumento da ocorrência de travagens bruscas leva também a um aumento do consumo esperado, pelo que se devem promover práticas de condução mais defensivas e menos agressivas, que implicam ganho no conforto para o passageiro mas também, de acordo com este resultado, ganhos ao nível do consumo.

Na Figura A 23, em anexo, apresenta-se o gráfico de elasticidade do consumo de combustível relativa à percentagem de ocorrência do evento 1007. Tal como relativamente à variável anterior, o consumo de combustível apresenta uma elasticidade consideravelmente baixa. Para uma variação da percentagem de ocorrência do evento 1007 de 0.9% para 2%, é esperada uma diminuição do

consumo na ordem dos 0.8%. Estes são valores novamente baixos, mas há que ter em conta que se referem a pequenas práticas de condução por parte dos motoristas, sem alteração de todos os outros factores que influenciam mais o consumo de combustível, pelo que mesmo que sejam pequenas as alterações devem ser registadas de forma a proceder-se à sua correcção nas sessões de formação de motoristas.

6.2.2 Validação do modelo de Motoristas

No modelo de motoristas conseguiu-se recolher uma nova amostra de validação com 530 motoristas, já eliminando os casos que não cumpriam os mesmos critérios definidos na amostra utilizada para o desenvolvimento do modelo de regressão. O ajustamento da amostra de validação ao modelo desenvolvido foi de 0.846, resultado que, apesar de ser inferior ao valor obtido para os dados originais ($R^2=0.916$), indica que existe uma aproximação ao modelo muito pronunciada, tal como se pode verificar através da observação da Figura 26. Este voltou a ser o modelo que apresentou a menor correlação com os dados os recolhidos, o que também era de esperar dado que não é muito comum que, com dados diferentes daqueles que criaram o modelo, se obtivessem valores de ajustamento superiores aos que se tinham verificado anteriormente para os dados originais. Observando a Figura 26, conclui-se que os casos se distribuem em torno da recta de maior tendência e que esta está muito próxima da recta de valores previstos pelo modelo de Motoristas.

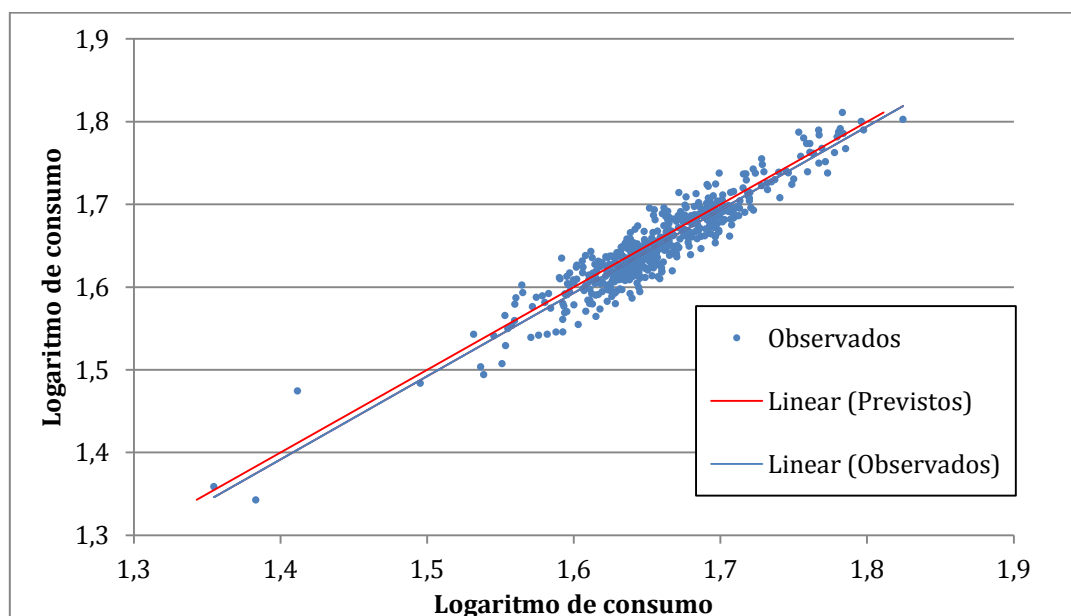


Figura 26 - Comparação entre valores previstos de logaritmo de consumo pelo modelo de Motoristas e valores observados da amostra de validação.

6.3 Modelo Viaturas

6.3.1 Considerações gerais

O modelo desenvolvido para a estimação do consumo das viaturas apresentou um valor de explicação da variância total da variável dependente de 95,4%, o que, tal como nos modelos anteriores, consiste num valor bastante significativo. Fazendo a validação do modelo com outra amostra de dados recolhida, conseguiu-se uma explicação da variância na ordem dos 90,2%, mantendo-se assim uma taxa de explicação bastante elevada da variável dependente.

O modelo de regressão obtido para as viaturas tem como variável dependente o logaritmo do consumo, uma vez que se verificou que esta transformação permitia uma melhor correlação com as variáveis independentes.

6.3.2 Análise de variáveis

6.3.2.1 Tipologia de veículos

Neste modelo só foi necessário introduzir uma variável para definir a tipologia de veículos, sendo essa a massa (neste caso transformada para o seu logaritmo, já que resultou num melhor ajustamento à variável dependente). Isto deve-se ao facto de, para cada tipologia, estar associado apenas um valor de massa distinto entre cada uma das tipologias.

Esta é a variável que apresenta o maior valor de coeficiente estandardizado, o que permite concluir que é muito relevante para o consumo esperado de uma viatura. Estes resultados eram, naturalmente, bastante espectáveis, tal como sugerido em SIMÕES, 2005, dadas as diferenças de consumo médio verificados entre os vários tipos de veículos. No entanto, os resultados vêm reforçar a utilidade da política de renovação de frota da RL, tendo em conta a substituição de veículos mais pesados por veículos mais leves.

Uma vez que no modelo de Viaturas a massa foi transformada no seu logaritmo, a elasticidade corresponde ao coeficiente de regressão obtido, que neste caso é 0.776, representando a maior elasticidade entre todas as variáveis nos 3 modelos.

Através da comparação com os resultados obtidos para os restantes modelos, é possível concluir que os 3 modelos apresentam ordens de valores de consumo para cada tipologia de veículo muito semelhantes, o que serve como validação entre os próprios modelos.

6.3.2.2 Velocidade comercial

A velocidade comercial média é, mais uma vez, uma variável independente que apresenta um alto valor (em absoluto) de coeficiente estandardizado. Também neste modelo volta a apresentar um sinal negativo, o que implica que o aumento desta variável independente se traduz numa expectável redução de consumo.

O gráfico de elasticidades (ver Figura A 24, em anexo) permite verificar que a elasticidade do consumo de combustível é ainda superior do que a verificada para os 2 modelos anteriores, sendo que o seu máximo valor situa-se numa elasticidade acima dos 0,2 o para variações da velocidade comercial.

A título de exemplo, é possível concluir que para um aumento da velocidade comercial de 18,3km/h para 21km/h (que constitui uma variação de 14,8%), pode-se esperar uma diminuição de 3,7% no consumo de combustível para uma viatura.

6.3.2.3 Comprimento médio de percurso

A variável relativa ao comprimento médio de percurso voltou também a ser seleccionada, desta vez para o modelo de viaturas. Apresentando um coeficiente estandardizado de -0.140, muito próximo do valor registado para o modelo de motoristas, significa que uma diminuição de percurso indicia a um aumento do consumo. Tal como se sugeriu anteriormente, este resultado deverá estar relacionado com o facto de as carreiras com percursos mais longos circularem a velocidades comerciais mais próximas do ideal, pelo que se obtém uma melhor eficiência do consumo de combustível.

A Figura A 25, em anexo, permite identificar a elasticidade que o comprimento médio de percurso induz no consumo médio de cada viatura, atingindo um máximo de elasticidade de 0.07.

A título de exemplo, é possível concluir que para um aumento da velocidade comercial de 5km/h para 10km/h, pode-se esperar uma diminuição de 3,6% no consumo de combustível da carreira em questão.

Fazendo uma análise à relação da variável comprimento médio de percurso com outras variáveis desenvolvidas, considerou-se interessante fazer uma comparação com a velocidade comercial. Através do gráfico de dispersão, que se apresenta em seguida, é possível detectar que existe uma forte relação positiva entre o comprimento médio de percurso com a velocidade comercial, o que significa que quanto maior for o comprimento médio de percurso maior a tendência para um aumento da velocidade comercial. Este resultado deverá ter um significado muito relacionado com a própria rede de carreiras da *Rodoviária de Lisboa*, e também com o facto de que algumas das suas carreiras de percurso mais longo circulem por troços com melhores condições ou até mesmo com

várias vias por sentido e com velocidades de circulação superiores. Devido a este facto também a distância entre paragens será superior o que, à partida, tenderá para um aumento da velocidade comercial.

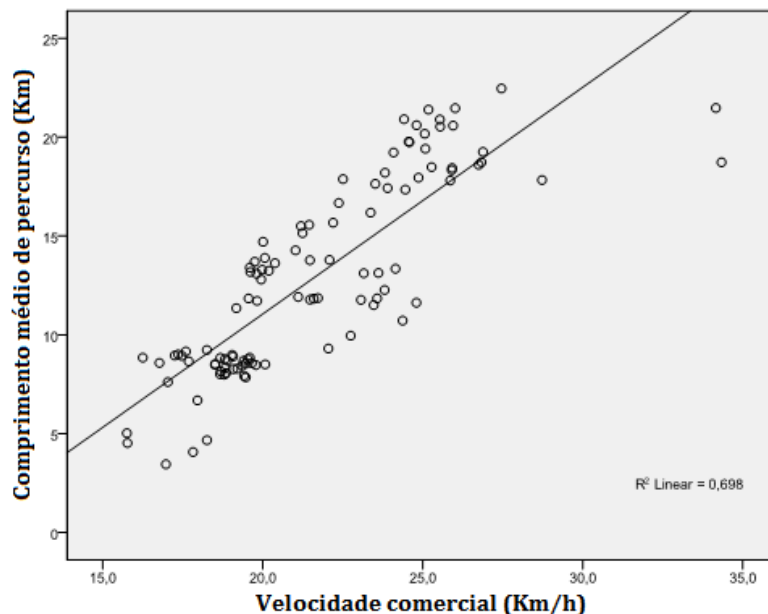


Figura 27 – Velocidade comercial vs Comprimento médio de percurso.

6.3.2.4 Percentagem de ocorrência do evento 1067 (as rotações acima do valor máximo estipulado)

O evento relativo à ocorrência rotações acima do valor máximo estipulado apresenta uma correlação positiva com o consumo, sendo que o seu coeficiente estandardizado é +0.077. Desta forma, este resultado sugere que o aumento da ocorrência de rotações acima do valor máximo estipulado leva também a um aumento do consumo esperado, pelo que se devem promover práticas de condução mais correctas.

Em anexo, na Figura A 26, apresenta-se o gráfico de elasticidade do consumo de combustível em relação à percentagem de ocorrência do evento 1067 - carregamento do ar em excesso de rotação. Pela observação do gráfico, facilmente se conclui que o consumo apresenta uma elasticidade pequena face a alterações da percentagem de ocorrência do evento 0167, mas ainda assim é superior à das outros eventos de condução seleccionados nos outros modelos.

A título exemplificativo, para uma variação da percentagem de ocorrência do evento 1067 de 0% para 0,3% há um aumento do consumo de combustível de 1,2%.

6.3.2.5 Idade média de motoristas

A variável idade média de motoristas também foi considerada como significativa na previsão do consumo, no modelo de viaturas. Contrariamente ao que tem sido apontado por alguns estudos (ÅBERG E RIMMÖ, 1998), este resultado indica que motoristas mais idosos tendem a ser menos receptivos à alteração de práticas erradas de condução, o que implica um menor aproveitamento das aulas de monitorização formativas. Este resultado chegou, foi corroborado por parte dos responsáveis da RL aquando da sua apresentação à empresa, pelo que está de acordo com a sua realidade.

A variável idade média de motoristas foi transformada no seu logaritmo e, portanto, a elasticidade do consumo de combustível corresponde ao coeficiente de regressão obtido, que neste caso foi 0.195, que é um coeficiente de elasticidade já mais elevado do que o apresentado pelas variáveis relativas à percentagem de ocorrência de eventos 1007 e ao comprimento médio de percurso. Ainda assim, uma variação da idade média de motoristas de 40 para 45 anos leva a um aumento do consumo de combustível de 2,4%.

Procurou-se detectar relações entre os eventos de condução e a idade dos motoristas, de forma a explicar o porquê da idade ser um factor aparentemente influente no consumo de combustível. No entanto, não foi possível detectar qualquer relação com os eventos recolhidos, como se demonstra, como exemplos, nos seguintes gráficos de dispersão

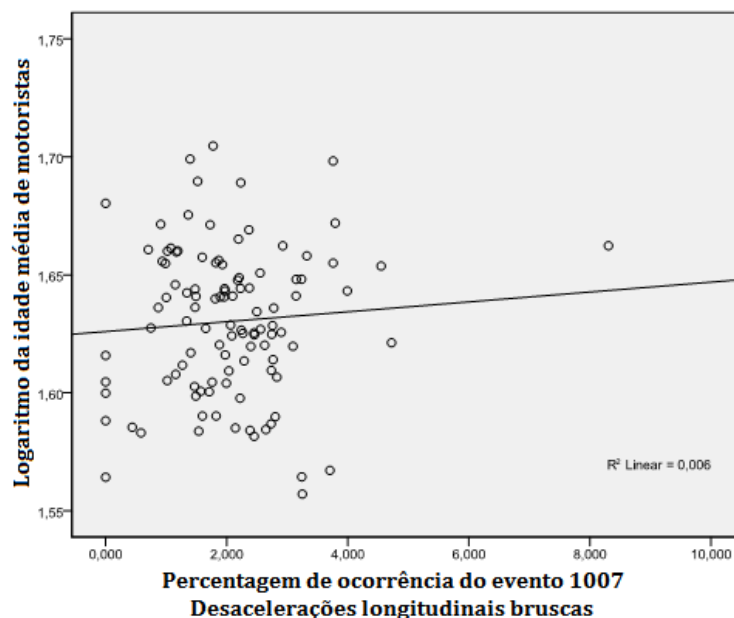


Figura 28 – Percentagem de ocorrência do evento 1007 vs Logaritmo da idade média de motoristas.

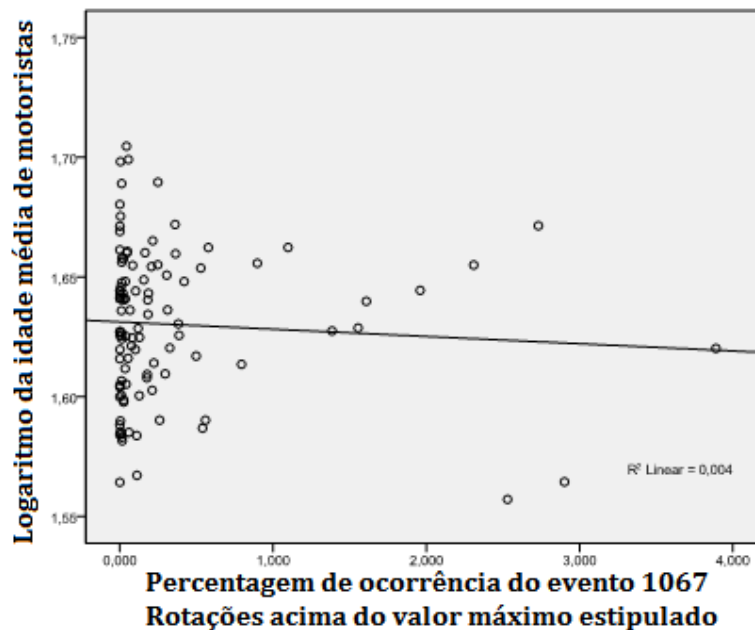


Figura 29 – Percentagem de ocorrência do evento 1007 vs Logaritmo da idade média de motoristas.

6.3.3 Validação do modelo de Viaturas

Por fim, relativamente ao modelo de viaturas, recolheu-se uma amostra de validação de 93 viaturas (a dimensão da amostra recolhida foi muito superior mas, mais uma vez, devido à inexistência de dados relativamente ao registo de eventos de condução em grande parte dos veículos, foram eliminados cerca de 2/3 dos casos recolhidos). Para o modelo de viaturas obteve-se, com a amostra de dados que o originou, um valor de explicação da variância de $R^2=0.952$. Tendo-se procedido à verificação do valor do coeficiente de determinação para a amostra de validação, o valor obtido foi de $R^2=0.902$, o que se traduz, ainda assim, num valor muito relevante de relação com o modelo desenvolvido, pelo que se pode validar a usabilidade do mesmo como previsão de consumo das viaturas. Através da observação da Figura 30, é interessante verificar que a recta de maior tendência dos valores observados se aproxima bastante da recta de valores previstos pelo modelo.

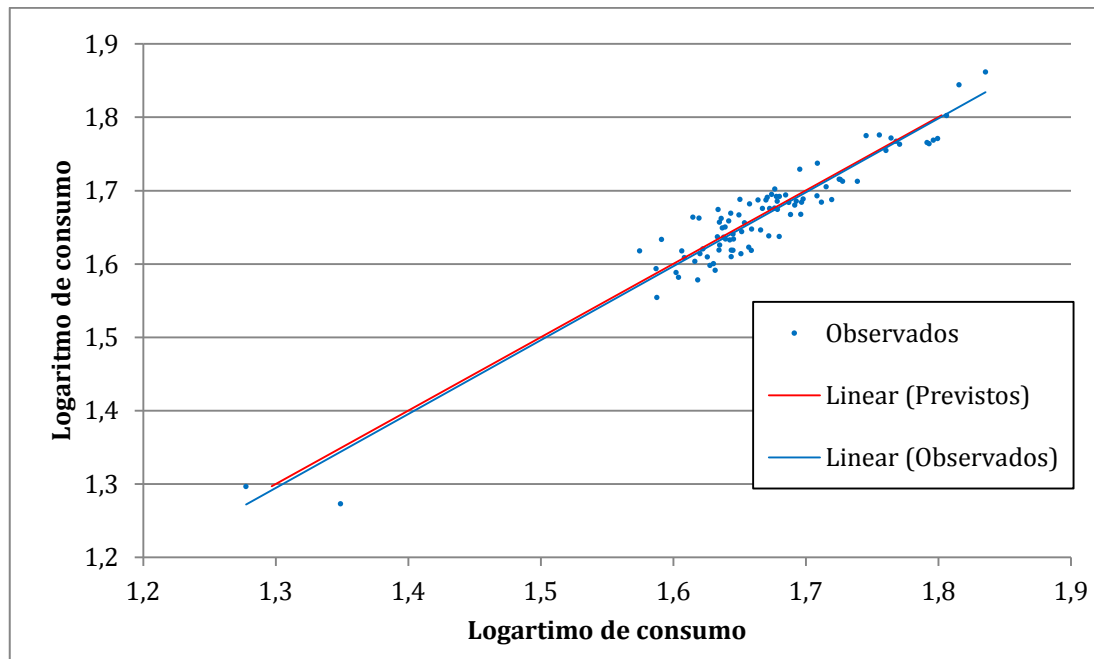


Figura 30 - Comparação entre valores previstos de logaritmo de consumo pelo modelo de Viaturas e valores observados na amostra de validação.

7. Conclusões

O transporte colectivo de passageiros assume um papel de grande importância na procura de uma mobilidade sustentável, uma vez que é o mais flexível e é aquele que melhor se adapta aos restantes modos de transporte, favorecendo a intermodalidade.

Actualmente, já existe uma grande gama da capacidade de transporte, desde o pequeno e económico mini-bus até aos autocarros articulados, com uma lotação de cerca de 200 pessoas, permitindo uma melhor adaptação à procura. Além disso, há várias formas de promover o aumento da velocidade comercial, seja a partir de vias dedicadas, sinalização prioritária, etc. o que permitirá o aumento quer da qualidade do serviço prestado quer da sua eficiência energética.

No entanto, os investimentos tanto ao nível da renovação de frota como dos sistemas de exploração acabam por ser bastante onerosos, pelo que ainda é necessário o apoio do Estado em operadoras de transporte rodoviário urbano.

Segundo os temas abordados no estado da arte, os factores detectados como mais influentes no consumo de combustível foram a massa do veículo (SIMÕES, 2005), a velocidade comercial (UITP, 2009), a topografia da rede viária (HENSHER, 2003) ou ainda o estilo de condução dos motoristas (VLIEGER, 2000). Todos estes resultados foram confirmados na presente dissertação, tendo sido seleccionados ainda outros factores influentes na previsão do consumo de combustível.

O presente trabalho teve como objectivo desenvolver modelos de regressão linear que permitissem prever o consumo de combustível dos três principais intervenientes de um serviço de transporte de passageiros: a carreira, o motorista e a viatura. O desenvolvimento desses modelos levou à obtenção de resultados bastante interessantes quer ao nível da sua fiabilidade quer das conclusões que deles se podem retirar, inseridos no âmbito de um processo de melhoramento da eficiência energética da *Rodoviária de Lisboa*, em que um com especial interesse foi o número de sessões de monitorização de motoristas.

Adicionalmente, os modelos obtidos para cada interveniente energético foram confrontados com os dados de uma recolha posterior e independente às amostras de dados que originaram cada modelo, comprovando-se uma aproximação muito forte à realidade por parte dos modelos de predição de consumo.

Para a realização deste trabalho foi recolhido um extenso conjunto de dados necessário à descrição de todos os intervenientes da actividade da *Rodoviária de Lisboa*. Com estes dados obtiveram-se três modelos de regressão que permitem a previsão do consumo de combustível para uma determinada carreira, motorista ou viatura. As amostras recolhidas permitiram a caracterização

completa de todas as variáveis definidas para 87 carreiras, 488 motoristas e 105 viaturas, nos meses de Outubro e Novembro de 2009 e Março de 2010.

7.1 Principais resultados

Através do desenvolvimento de modelos de regressão linear, foi possível recolher informação sobre os factores que influenciam mais o consumo de combustível para três modelos relativos a carreiras, motoristas e viaturas.

Os resultados apresentam uma boa significância estatística mas, além disso, é possível identificar de que forma o consumo de combustível é afectado, pelo que se apresentam, em seguida, algumas conclusões sobre estes resultados, referindo os principais agentes detectados como influentes para esta análise.

7.1.1 Tipologia de viaturas

As variáveis relativas à tipologia de viaturas revelaram-se como sendo aquelas que maior influência têm na previsão do consumo. Foram utilizadas variáveis relativas à percentagem de utilização de cada tipo de viatura ou uma variável relativa à respectiva massa, sendo que, em qualquer dos casos, foram sempre estas as variáveis independentes que mais se destacaram das restantes nos três modelos desenvolvidos. Este resultado vem confirmar a importância do programa de renovação de frota já posto em prática pela *Rodoviária de Lisboa*, com vista à adição na sua frota de viaturas mais leves, de tipologias Mini e Midi em detrimento de viaturas Articuladas, renovação também associada a uma diminuição da procura e consequentes ajustes de oferta. De acordo com o modelo de carreiras, a troca entre uma viatura Articulada por um Mini-bus representa uma redução de consumo na ordem dos 70%.

7.1.2 Velocidade comercial

A velocidade comercial foi a variável independente que, além das variáveis de tipologia de viatura, foi seleccionada em todos os modelos de regressão. Isto vem comprovar a sua importância no consumo de combustível, pelo que é importante que se faça uma análise directa e contínua à velocidade comercial verificada, nomeadamente em cada carreira, uma vez que além da melhoria na eficiência energética este também é um aspecto muito importante em relação à satisfação do cliente. Assim, é importante promover medidas que suportem um aumento da velocidade comercial, pois disso dependerá o sucesso de um serviço em todos os aspectos.

7.1.3 Monitorizações formativas de motoristas

Outra das variáveis detectadas como influentes foi o número de monitorizações médio de motoristas no modelo de carreiras. Este resultado é bastante interessante e surge como uma confirmação do sucesso desta iniciativa desenvolvida pela *Rodoviária de Lisboa*, uma vez que está associado ao número médio de monitorizações uma diminuição do consumo de combustível. Para uma variação do número médio de monitorizações por carreira de 2,5 para 4 sessões o modelo de carreiras prevê uma diminuição de 4% do consumo de combustível. Assim sendo, serve este resultado também como incentivo ao prolongamento desta prática que, através do estudo desenvolvido neste trabalho, se confirma como uma iniciativa de sucesso.

7.1.4 Idade motoristas

A idade de motoristas foi seleccionada como um factor que influencia o consumo de combustível. Este é um resultado que já era do conhecimento da *Rodoviária de Lisboa*, pelo que serviu como confirmação do mesmo. A sua explicação, em princípio, prende-se com o facto de motoristas mais experientes estarem menos disponíveis para alteração das suas práticas erradas de condução, nomeadamente quando são alertados para as mesmas nas sessões de monitorização, o que levará a que a eficiência energética dos seus serviços seja mais reduzida.

7.1.5 Eventos de condução

Uma das partes deste trabalho com maior interesse consistia em avaliar a influência que os eventos de condução dos motoristas registados teriam sobre o consumo de combustível. Nesse sentido, foram detectados 3 eventos de condução que sugerem ter uma influência negativa sobre o consumo de viaturas.

Através do modelo de carreiras foi detectada uma relação entre as desacelerações longitudinais bruscas e o consumo de combustível, sugerindo uma correlação positiva. De acordo com o modelo de Carreiras é espectável que uma diminuição de 0,04% para 0,00% da percentagem de ocorrência deste evento leve a uma diminuição do consumo de combustível de 0,4%.

Outro evento de condução detectado como potenciador do aumento do consumo de combustível foi o carregamento do ar em excesso de rotação, sendo que, para a anulação (0%) da percentagem de ocorrência deste evento a partir de um caso que verificasse 2.00% dessa mesma percentagem de ocorrência é esperada uma diminuição no consumo de combustível na ordem dos 1.50%.

Por fim, no modelo de viaturas foi seleccionada a variável rotações acima do valor máximo estipulado. De acordo com o modelo, uma pequena de 0.059% para 0.3% da percentagem de ocorrência do evento 1067 implica uma variação de 1.0% no consumo de combustível.

Os resultados obtidos provam-se sensíveis às variações da percentagem de ocorrência dos eventos, pelo que devem ser aproveitados como um estímulo nas monitorizações formativas de forma a alertar os motoristas para estes erros de condução, uma vez que se repercutem sobre o combustível dispendido no serviço de transporte. Além disso, quer as travagens bruscas quer o excesso de rotações podem indiciar aspectos de uma condução menos confortável para o passageiro.

7.2 Outras medidas de conservação de energia

É ainda necessário referir medidas de intervenção que seriam vantajosas ao desempenho energético da RL, mas que se considera serem de mais difícil implementação ou em relação às quais a empresa não terá grande poder de execução:

- Melhorar as condições de circulação através da redução de veículos mal estacionados e da implementação de pistas BUS em locais urbanos de difícil circulação, de modo a aumentar a velocidade comercial;
- Implementar sistemas semafóricos prioritários para o transporte público colectivo;
- Evitar a colocação de paragens de autocarro em troços dos percursos com inclinações mais elevadas.

7.3 Limitações de metodologia desenvolvida

Ao longo do processo de recolha de dados e de desenvolvimento dos modelos de regressão foram detectadas limitações desfavoráveis à obtenção de resultados ainda mais precisos ou que impediram a verificação de certos efeitos que se pretendiam testar inicialmente. Assim, em seguida serão descritas algumas das situações que deveriam ser alteradas ou corrigidas de forma a que, no futuro, se desenvolvam análises que possam registar novos indicadores na predição de consumo de combustível.

O consumo médio obtido diariamente é calculado através do quociente entre a quantidade de combustível com que se abastece uma viatura no fim da sua actividade diária e os quilómetros percorridos pela mesma. Este método revela-se pouco preciso, quer devido ao facto de o enchimento do depósito não constituir uma medida exacta, quer devido aos erros humanos aliados à anotação das quantidades de combustível com que se abastecem as viaturas (problema detectado na base de dados recolhida). Uma vez que o combustível implica uma grande fatia dos gastos da

empresa, considera-se muito importante que se optem por técnicas de controlo do mesmo mais rigorosas de forma a que os resultados das análises que se vierem a desenvolver também o possam ser.

No desenvolvimento de variáveis relativas à inclinação de cada carreira, optou-se por criar variáveis que considerassem intervalos de inclinação, em módulo (exemplificando, criou-se uma variável para inclinações superiores a 5% e outra com inclinações inferiores a 5%) independentemente do sentido do movimento (ascendente ou descendente). No futuro, para análises mais detalhadas mas menos abrangentes, por exemplo a uma determinada carreira, poderá ser realizado o estudo apenas num dos sentidos dessa carreira de forma a conseguir-se distinguir a relação que têm no consumo de combustível as inclinações positivas e negativas de uma forma mais exacta.

Actualmente, não é possível nas viaturas da *Rodoviária de Lisboa* o registo do número de passageiros que circulam em cada troço de uma carreira. Tendo em conta a influência da massa de uma viatura no seu consumo de combustível, considera-se muito importante que sejam desenvolvidos esforços no sentido de poder obter esta informação. Uma vez que qualquer metodologia associada à recolha de informação possa ser complicada, sugere-se que se façam primeiramente análises a um número reduzido de viaturas para se poder obter resultados que possam verificar a utilidade desta informação, que poderia ser importante na definição do tipo de viaturas a associar a um determinado serviço de transporte de passageiros.

Na amostra do modelo de viaturas não foram utilizados casos pertencentes à tipologia Midi. Este problema poderá inviabilizar a utilização do modelo desenvolvido para estimação do consumo de combustível esperado em viaturas desta tipologia. Assim, considera-se que existe uma falha de informação que deve ser corrigida por parte da *Rodoviária de Lisboa*.

Os resultados obtidos nestes modelos têm uma forte relação com a realidade vivida nesta empresa, não sendo portanto recomendável a sua utilização em situações diferentes daquelas para quais foram calculados. Para tal, deverá proceder-se a uma recolha de dados relativa ao novo caso de estudo.

7.4 Trabalhos futuros

Com o objectivo de continuar a evoluir no sentido de melhorar a eficiência energética do serviço prestado pela Rodoviária de Lisboa, é importante sugerir formas de intervenção e análise de desempenho sobre a sua actividade. Nesse sentido, em seguida indicam-se as propostas que se consideram relevantes neste processo:

- Deve ser feito um esforço de forma a melhorar as condições de recolha de dados relativos aos vários intervenientes, nomeadamente em relação às limitações de recolha apontadas;
- Através dos modelos desenvolvidos, deve ser feita a detecção de casos que diferem muito dos resultados esperados pelos modelos e qual a razão desse afastamento (deve ter-se em conta que esta metodologia poderá carecer de uma adaptação constante à realidade);
- Deverá ser feita a avaliação da relação benefício/custo das várias medidas de conservação de energia;
- Poderá ser interessante definir modelos similares relativos a cada centro de actividade de transporte da Rodoviária de Lisboa, uma vez que cada CAT apresenta condições de tráfego e condições topográficas muito distintas;
- Devem-se traçar objectivos relativamente às monitorizações formativas de motoristas, de forma a estimular o desenvolvimento dos mesmos e a diminuição do consumo de combustível;
- Por fim, deve-se fazer a validação experimental periódica dos resultados obtidos para cada modelo e, se necessário, reajustá-los.

Referências bibliográficas

- ÅBERG, L., & Rimmö, P.-A. (1998), *Dimensions of aberrant driver behavior*. *Ergonomics*, 41(1), 39-56.
- ADA (2009), Apontamentos da cadeira de Análise de Dados e Avaliação, do ano lectivo de 2009/2010.
- ANDRÉ, José Manuel C. S. (2008), *Transporte Interurbano em Portugal – O sistema actual e os seus desafios*, Lisboa, IST Press, Março.
- BAICHUAN, Lu e Viegas, José Manuel (2001), *Sistemas de controlo de tráfego com faixas bus intermitentes*, Dissertação de Doutoramento em Engenharia Civil no Instituto Superior Técnico, Lisboa, 10 de Dezembro.
- BARNIOL, R. Warlam (2005), *Bus Network Operations and Management in Chicago and Barcelona: The effects of larger stop spacing on social profitability*, Barcelona
- BORREGO, Carlos (2005) Seminário: *Transporte sustentável em zonas urbanas*, Universidade de Aveiro, Departamento Ambiente e Ordenamento, 12 de Janeiro.
- CAMPBELL, K. L. (1991). Fatal accident involvement rates by driver age for large trucks. *Accident Analysis and Prevention*, 23, 287-295.
- CE (2001) LIVRO BRANCO: *A Política Europeia de Transportes no Horizonte 2010: a Hora das Opções*, Luxemburgo, Serviço das Publicações Oficiais das Comunidades Europeia.
- CE (2002), *Energy and Transport in figures*, Energy and Transport DG in co-operation with EUROSTAT.
- CE (2003a), *A Europa numa Encruzilhada – A necessidade de transportes sustentáveis* Comissão Europeia Direcção-Geral da Imprensa e Comunicação, Bruxelas Junho de 2003.
- CE (2003b), *European Energy and Transport Trends to 2030*, Direcção-Geral da Energia e Transportes, Janeiro
- CE (2006a), *Directiva 2006/67/CE do Conselho, de 24 de Julho de 2006, que obriga os Estados-Membros a manterem um nível mínimo de reservas de petróleo bruto e/ou de produtos petrolíferos*.
- CE (2006b), *Road Transport Policy – Open roads across Europe*, Serviço das Publicações Oficiais das Comunidades Europeia, 15 de Novembro.
- CE (2007), *LIVRO VERDE Por uma nova cultura de mobilidade urbana*, COM(2007) 551 final, Bruxelas, 25 de Junho.

CE (2009), *COMUNICAÇÃO DA COMISSÃO: Um futuro sustentável para os transportes: rumo a um sistema integrado, baseado na tecnologia e de fácil utilização*, COM (2009) 279 final; Bruxelas, 17 de Junho.

CHARNET, Reinaldo e outros (1999), *Análise de modelos de regressão linear com aplicações*, Editora Unicamp, Campinas.

CHATERJEE, Samprit e HADI, Ali S. (2006) *Regression Analysis by example*, 4ª Edição, Wiley-Interscience, Nova Jérícia

CHIPMAN, M. L., MacGregor, C. G., Smiley, A. M., Lee-Gosselin, M. (1993). *The role of exposure in comparisons of crash risk among different drivers and driving environments*. *Accident Analysis and Prevention*, 25, 207-211.

COHEN, Jacob e outros (2003), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3ª Edição, LEA, Nova Jersey.

CUBUKCU, K. Mert e Eylul, Dokuz (2006) *Cost of Urban Bus Transit Operations and Geography of Service Territory*, Bruxelas, 10 de Março.

EEA (2001), *External costs of transport*, European Environment Agency, 20 de Agosto.

EUROSTAT (2009), *Panorama of Transport*, Eurostat Statistical books, Luxemburgo.

FARIAS, Tiago (2009), *O que é a Eco-condução: Critérios e Importância*, Seminário Eco-condução: Economia, Segurança e Ambiente, Lisboa, 20 de Maio.

GARAVAGLIA, S., Sharma, A. (2009), *A smart guide to dummy variables: four applications and a macro*, Dun & Bradstreet, Murray Hill, Nova Jérícia.

GAUR, Ajai S. e Gaur, Sanjaya S. (2009) *Statistical Methods for Practice and Research*, 2ª Edição, Response Books, Nova Dehli.

GV (2006), *Posição escrita do Grupo dos Verdes/Aliança Livre Europeia aprovada pelo Grupo*, 13 de Dezembro.

HAIR, Joseph F. e outros (1995), *Multivariate Data Analysis: With Readings*, 4ª Edição, Prentice Hall College Div.

HENSHER (2003) *Influences on distance-based unit costs for non-urban bus operations in New South Wales*. *Road and Transport Research* 12(1): 16-27.

HO, Robert (2006), *Handbook of univariate and multivariate data analysis and interpretation with SPSS*, Chapman & Hall /CRC, Rockhampton.

IMTT (2008), *Annual Review of Energy Policies*, IMTT, Novembro de 2008.

KENWORTH (2008), *White paper on fuel economy*, Agosto.

KOSHAL, R. K. (1970), *Economics of Scale in Bus Transport: Some United States Experience*, Journal of Transport Economics and Policy, 4, pp. 29-36.

KYOUNGHO, Ahn, e outros (2002), *Estimating Vehicle Fuel Consumption and Emissions Based on Instantaneous Speed and Acceleration Levels*, Journal of Transportation Engineering, Vol. 128, No. 2, March/April 2002, pp. 182-190

LANDAU, Sabine, Everitt, Brian S. (2004), *A Handbook of Statistical Analyses using SPSS*, Chapman & Hall/CRC Press LLC, Florida.

LYNCH, K (1971), *Site Planning*, Cambridge, The MIT Press.

MAROCO, João (2007), *Análise estatística com utilização do SPSS*, 3ª Edição, Edições Sílabo, Lisboa.

MARSH WM (1983) *Landscape Planning: Environmental Applications*. Reading, MA: Addison-Wesley Publishing Company.

MOURA, F., Figueira, P., Vieira, J. (2003), *Programa Nacional para as Alterações Climáticas, Estudos Base. Cenário de Referência do Sector dos Transportes*. CESUR – Centro de Sistemas Regionais e Urbanos.

MOURA, Lara Cordes Lima Trigueiro de (2007), *Análise energética e ambiental da aplicação de biocombustíveis em frotas urbanas*, Dissertação para a obtenção do Grau de Mestre em Hidráulica e Recursos Hídricos, Instituto Superior Técnico, Lisboa.

NETER, J.; Wasserman, W. e Kutner, M. H. (1985), *Regression, analysis of variances and experimental designs*. 2ª.Edição, Richard D. Irwin.

PALLANT, Julie (2007), *SPSS Survival Manual*, 3ª Edição, McGraw Hill, Sydney.

Petroleum Conservation Research Association (PCRA) (2005) *Tested Approaches to Save Diesel in Buses and Trucks* (on-line publication). Acedido em: Agosto de 2010, em: <http://www.pcr.org/transport/saveDiesel.htm>.

PNAC (2006) Instituto do Ambiente; *Plano Nacional para as Alterações Climáticas, Anexo Técnico – Transportes*, Presidência do conselho de ministros, 23 de Agosto.

PROPLANO (2002), *Estudo sobre o sistema de transportes colectivos de passageiros da área metropolitana de Lisboa*, Proplano, Lisboa, Dezembro

RAKHA, Hesham e Ding, Yonglian (2003), *Impact of Stops on Vehicle Fuel Consumption and Emissions*, ASCE - JOURNAL OF TRANSPORTATION ENGINEERING, 2003 Vol.129 No.1 Janeiro / Fevereiro.

REA (2007), *Relatório do Estado do Ambiente 2006*, Outubro.

- RL (2008a), *Plano de Racionalização Energética da Rodoviária de Lisboa, S.A. Relatório de Progresso anual*, Lisboa, 8 de Setembro.
- RL (2008b) *GISFROT: Programa de Melhoria da Qualificação dos Motoristas e da Qualidade de Serviço*
- RL (2009), *Manual GisFrot*, Rodoviária de Lisboa S.A., Lisboa.
- SHORT, J. (2003), *Foreword*, Fifty years of transport policy, European Conference of Ministers of Transport, pp 3-4, OECD Publication Services, Paris.
- SIMÕES, A. (2005), *Metodologia para Auditorias Energéticas a Frotas de Autocarros Urbanos*, Dissertação para a obtenção do Grau de Mestre em Transportes, Instituto Superior Técnico, Lisboa.
- T&E (2004), *Sense and Sustainability: Smart thinking to restart European transport policy*, and Stichting Natuur en Milieu, Bruxelas.
- TABACHNICK, Barbara G. e Fidell, Linda S. (2007), *Using Multivariate Statistics*, 5ª Edição, Pearson Education Inc., Boston.
- UITP (2009), *Bus Systems: An efficient mode of transport*, UITP's World Congress and Exhibition, Viena.
- VIEGAS, J.(2003), *Transport Policy Development in Europe 1950-2020*, Fifty years of transport policy, European Conference of Ministers of Transport, pp 105-115, OECD Publication Services, Paris.
- VIEIRA, João F. C. C. (2005), *Aplicação de instrumentos com objectivos ambientais no sector dos transportes: da intervenção regulamentar à multi-instrumentalidade*, Dissertação para a obtenção do Grau de Mestre em Transportes, Instituto Superior Técnico, Lisboa, Fevereiro.
- VIEIRA, Manuel A., (2004), *Melhoria da Velocidade dos Transportes Públicos de Superfície em Lisboa por Regulação da Admissão de Trânsito*, , Dissertação para a obtenção do Grau de Mestre em Engenharia Electrotécnica e de Computadores, Instituto Superior Técnico, Lisboa, Janeiro.
- VLIEGER, Keukeleere e Kretzchmar, J. (2000), *Environmental effects of driving behaviour and congestion related to passenger cars*, *Atmospheric Environment* (2000) pp. 4649-4655.
- VOLVO, *Consumo de combustível*, Acedido em 20 de Julho de 2010, em: http://www.volvotrucks.com/trucks/portugal-market/pt-pt/trucks/environment/pages/fuel_consumption.aspx.
- WASHINGTON, Simon P. e outros (2003), *Statistical and Econometric Methods for Transportation Data Analysis*, Chapman & Hall/CRC, Albuquerque.
- WILLIAMS, A. F. (1985). *Nighttime driving and fatal crash involvement of teenagers. Accident Analysis and Prevention*, 17, 1-5.

WUNSCH, P (1996) *Cost and productivity of major urban transit systems in Europe: An exploratory analysis*. The Journal of Transport Economics and Policy 30(2): 171-186.

ZARKADOULA, Maria e outros (2007), *Training urban bus drivers to promote smart driving: A note on a Greek eco-driving pilot program*, Department of Environment and Transport, Division of Energy Efficiency, Centre for Renewable Energy Sources, 19th km Marathonos Avenue, Attiki.

ANEXOS

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA**

**ANEXO I – TABELAS DESCRITIVAS ESTATÍSTICAS DE
VARIÁVEIS RELATIVAS A CARREIRAS**

ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA

Tabela A 1 - Sumário descritivo de variáveis relativas a Carreiras

		Duração média de percurso	Comprimento de percurso (m)	Número de paragens	Distância média entre paragens (m)	Distância mínima entre paragens (m)	Distância máxima entre paragens (m)	Inclinação = 0%	Inclinação < 5%	Inclinação > 5%	Inclinação > 8%	Velocidade Comercial	Tipo Carreira Urbana (%)	Tipo Carreira Suburbana (%)	Tipo Carreira Directa (%)	
N	Valid	87	87	87	87	87	87	87	87	87	87	87	87	87	87	
	Média	31,878	12383,95	31,72	417,54	130,45	1731,08	,975	70,8897	28,1333	10,4793	20,782	,14	,75	,11	
	Mediana	30,300	10277,00	29,00	349,00	118,00	942,00	,800	73,5000	26,2000	8,1000	19,500	,00	1,00	,00	
	Moda	15,3 ^a	2810 ^a	28 ^a	281 ^a	88 ^a	804 ^a	,0	50,20 ^a	8,40 ^a	2,70 ^a	19,4	0	1	0	
	Desvio padrão	11,7247	6727,355	14,196	310,595	76,744	2436,138	1,0816	12,56845	12,59312	6,82799	4,7852	,347	,437	,321	
	Variância	137,469	4,526E7	201,528	96469,55	5889,622	5934768,191	1,170	157,966	158,587	46,621	22,898	,120	,191	,103	
	Range	53,4	33264	72	2403	662	14685	5,7	54,70	54,80	33,30	24,1	1	1	1	
	Mínimo	10,9	2810	8	212	62	354	,0	38,50	6,70	1,70	13,5	0	0	0	
	Máximo	64,3	36074	80	2615	724	15039	5,7	93,20	61,50	35,00	37,6	1	1	1	
	Percentis	25	23,900	8283,00	22,00	296,00	96,00	761,00	,000	61,2000	19,2000	5,5000	17,300	,00	,00	,00
		50	30,300	10277,00	29,00	349,00	118,00	942,00	,800	73,5000	26,2000	8,1000	19,500	,00	1,00	,00
		75	37,600	15659,00	38,00	427,00	141,00	1520,00	1,400	79,7000	37,5000	15,5000	23,400	,00	1,00	,00

a. Existem múltiplos valores de moda. É apresentado o valor mais baixo.

ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA

Tabela A 2 - Sumário estatístico de variáveis relativas a Motoristas

	Anos trabalho	Monitorizações	Idade (anos)	Ev0120 (%)	Ev0121 (%)	Ev0130 (%)	Ev0140 (%)	Ev0141 (%)	Ev1005 (%)	Ev1007 (%)	Ev1009 (%)	Ev1010 (%)	Ev1014 (%)	Ev1015 (%)	Ev1026 (%)	Ev1058 (%)	Ev1060 (%)	Ev1062 (%)	Ev1067 (%)	Ev1069 (%)	Ev1070 (%)	
N	Valid 488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	488	
Média	10,402	3,5676	42,908	,00609	,47459	,00000	,02392	,82410	2,48981	2,01435	,06732	,47417	,23221	2,20470	,52207	4,16480	1,21456	,09334	,25592	,00118	2,48477	
Mediana	9,0000	3,0000	43,000	,00000	,04080	,00000	,00000	,05270	1,97900	1,69175	,04220	,08115	,02550	1,78470	,47835	2,65870	,58085	,00000	,10065	,00000	2,15080	
Moda	2,00	3,00	41,00	,000	,000	,000	,000	,000	,000 ^a	,000 ^a	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	
Desvio padrão	8,65910	1,65616	9,38574	,033788	6,391508	,000000	,208375	4,419094	1,882279	1,415703	,079495	1,16852	,901858	1,653521	,397055	3,815216	1,710399	,385230	,505806	,007460	1,865521	
Variância	74,980	2,743	88,092	,001	40,851	,000	,043	19,528	3,543	2,004	,006	1,365	,813	2,734	,158	14,556	2,925	,148	,256	,000	3,480	
Range	45,00	9,00	41,00	,704	100,000	,000	4,536	78,506	12,152	9,903	,707	12,365	11,437	11,763	7,042	25,458	17,222	4,706	5,821	,120	11,260	
Mínimo	,00	,00	22,00	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	
Máximo	45,00	9,00	63,00	,704	100,000	,000	4,536	78,506	12,152	9,903	,707	12,365	11,437	11,763	7,042	25,458	17,222	4,706	5,821	,120	11,260	
Percentis	25	3,0000	2,0000	35,250	,00000	,01855	,00000	,00000	,01163	1,10528	1,00148	,01945	,01285	,00000	,94780	,31088	1,76105	,24213	,00000	,02810	,00000	,98795
	50	9,0000	3,0000	43,000	,00000	,04080	,00000	,00000	,05270	1,97900	1,69175	,04220	,08115	,02550	1,78470	,47835	2,65870	,58085	,00000	,10065	,00000	2,15080
	75	14,000	5,0000	49,000	,00458	,08888	,00000	,00838	,26468	3,54665	2,63657	,08913	,39793	,08540	3,01448	,66465	5,18388	1,51828	,02692	,25915	,00000	3,63253

a. Existem múltiplos valores de moda. É apresentado o valor mais baixo.

Tabela A 3 - Sumário estatístico de variáveis relativas a Viaturas

		Tipologia Mini (%)	Tipologia Midi (%)	Tipologia Standard (%)	Tipologia Articulado (%)	Consumo (L/100km)	Massa (ton)	Idade (anos)
N	Valid	105	105	105	105	105	105	105
Média		,06	,00	,78	,16	47,080	15852,38	13,971
Mediana		,00	,00	1,00	,00	46,193	15600,00	16,000
Moda		0	0	1	0	19,1 ^a	15600	16,0
Desvio Padrão		,233	,000	,416	,370	10,3902	3318,077	4,5138
Variância		,054	,000	,173	,137	107,956	1,101E7	20,374
Range		1	0	1	1	58,3	15900	20,0
Mínimo		0	0	0	0	19,1	5000	1,0
Máximo		1	0	1	1	77,4	20900	21,0
Percentis	25	,00	,00	1,00	,00	42,280	15600,00	13,500
	50	,00	,00	1,00	,00	46,193	15600,00	16,000
	75	,00	,00	1,00	,00	51,040	15600,00	16,000

a. Existem múltiplos valores de moda. É apresentado o valor mais baixo.

**ANEXO II – OUTPUT DOS MODELOS DE REGRESSÃO
LINEAR MÚLTIPLA**

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA**

**Tabela A 4 - Sumário do modelo de regressão de Carreiras
Model Summary^b**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Carreiras	,986 ^a	,972	,970	,01554	1,682

a. Predictors: (Constant), TipMidi, TipMini, Dsup5_2, TipArtic, Ev0140, Distmax_log, Monitorizacoes, VelCom

b. Dependent Variable: Consumo_Log

**Tabela A 5 - Análise ANOVA do modelo de Carreiras
ANOVA^b**

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	,662	8	,083	342,747	,000 ^a
Carreiras Residual	,019	78	,000		
Total	,681	86			

a. Predictors: (Constant), TipMidi, TipMini, Dsup5_2, TipArtic, Ev0140, Distmax_log, Monitorizacoes, VelCom

b. Dependent Variable: Consumo_Log

**Tabela A 6 - Coeficientes de regressão do modelo de Carreiras
Coefficients^a**

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
Constante	1,741	,024		71,494	,000		
Número de monitorizações	-,012	,003	-,117	-4,201	,000	,458	2,183
Velocidade Comercial	-,005	,001	-,263	-7,788	,000	,311	3,220
Tipologia Mini	-,004	,000	-,888	-42,120	,000	,798	1,253
Quadrado da percentagem de inclinação superior a 5%	1,173E-5	,000	,104	3,713	,000	,450	2,224
Logaritmo da distância máxima entre paragens	,019	,008	,070	2,269	,026	,368	2,719
Percentagem de ocorrência do evento 0140	,048	,021	,049	2,308	,024	,795	1,257
Tipologia Articulada	,002	,000	,238	12,053	,000	,911	1,098
Tipologia Midi	-,001	,000	-,167	-8,565	,000	,933	1,071

a. Dependent Variable: Consumo_Log

**Tabela A 7 - Diagnóstico de resíduos do modelo de Carreiras
Residuals Statistics^a**

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1,3455	1,7862	1,6418	,08776	87
Std. Predicted Value	-3,376	1,645	,000	1,000	87
Standard Error of Predicted Value	,003	,015	,005	,002	87
Adjusted Predicted Value	1,3379	1,7932	1,6419	,08772	87
Residual	-,03321	,02890	,00000	,01480	87
Std. Residual	-2,137	1,859	,000	,952	87
Stud. Residual	-2,196	2,057	,000	1,003	87
Deleted Residual	-,03507	,03647	-,00004	,01649	87
Stud. Deleted Residual	-2,252	2,102	,001	1,012	87
Mahal. Distance	1,341	76,568	7,908	10,312	87
Cook's Distance	,000	,142	,013	,023	87
Centered Leverage Value	,016	,890	,092	,120	87

a. Dependent Variable: Consumo_Log

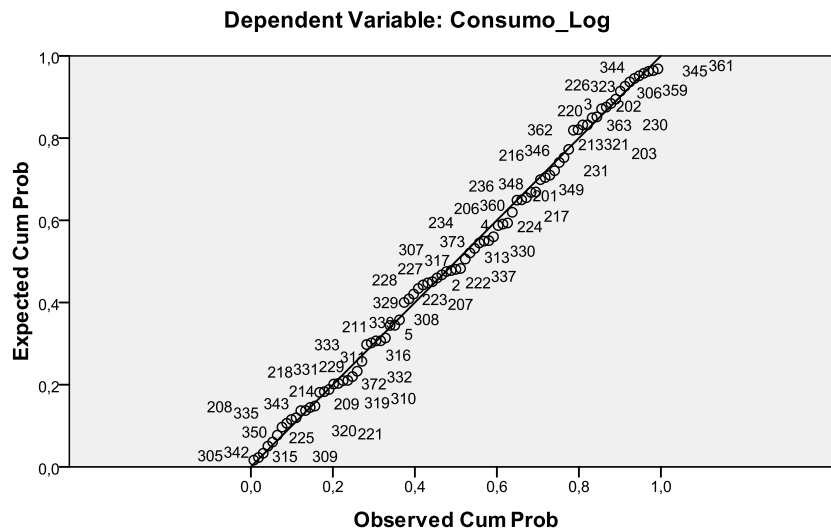


Figura A 1 - Gráfico de probabilidade normal dos resíduos do modelo de Carreiras.

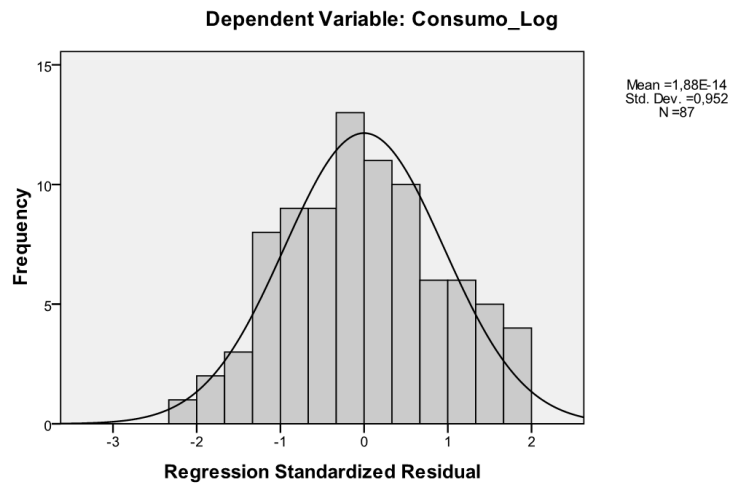


Figura A 2 - Histograma de resíduos do modelo de Carreiras.

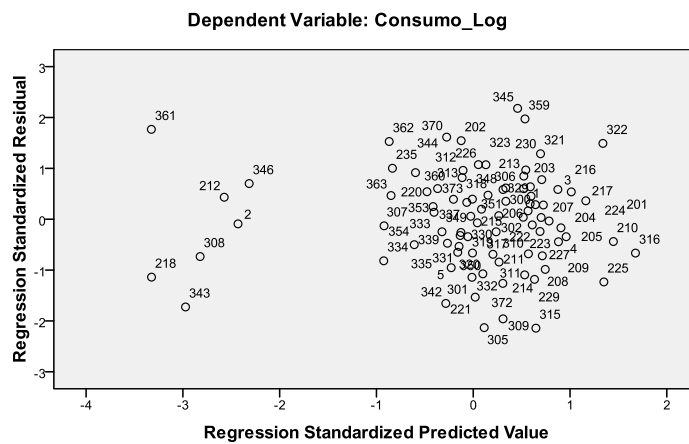


Figura A 3 - Gráfico de resíduos standardizados vs resíduos standardizados previstos do modelo de Carreiras.

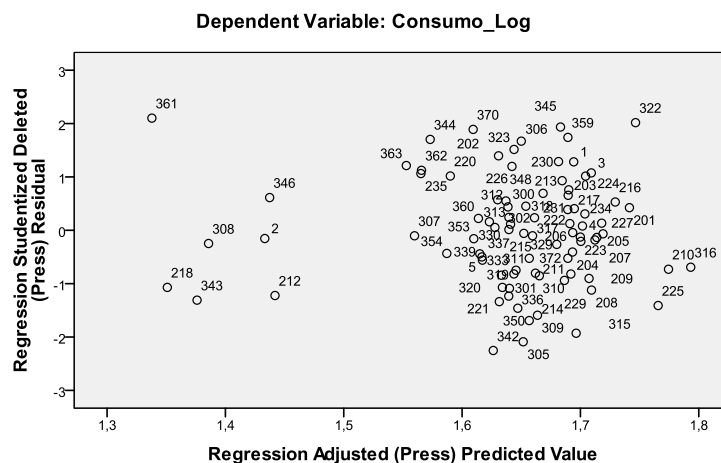


Figura A 4 - Gráfico de resíduos studentizados vs valores ajustados previstos (press) do modelo de Carreiras.

Tabela A 8 - Teste de normalidade de resíduos Kolmogorov-Smirnov do modelo de Carreiras

		Unstandardized Residual
N		87
Normal Parameters^{a,b}	Mean	,0000000
	Std. Deviation	,01480208
	Absolute	,049
Most Extreme Differences	Positive	,044
	Negative	-,049
Kolmogorov-Smirnov Z		,460
Asymp. Sig. (2-tailed)		,984
Exact Sig. (2-tailed)		,977
Point Probability		,000

a. Test distribution is Normal.

b. Calculated from data.

Tabela A 9 - Teste de normalidade de resíduos Shapiro-Wilk do modelo de Carreiras

Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
Unstandardized Residual	,986	87	,502

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

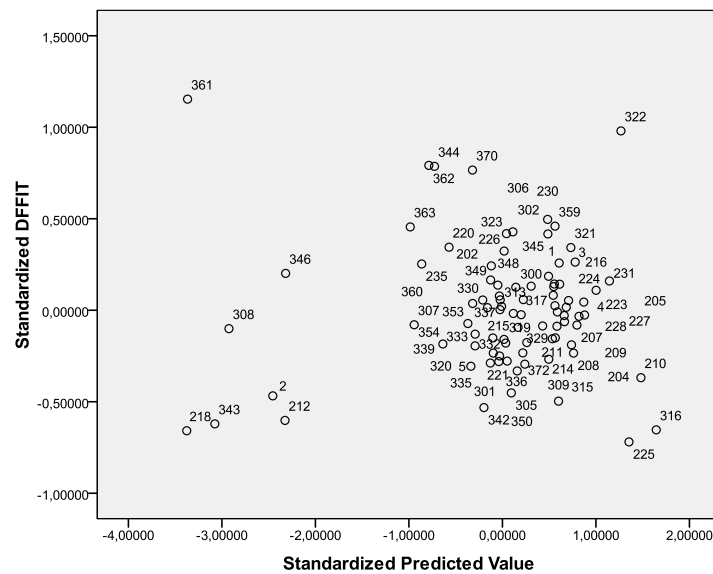


Figura A 5 - Gráfico de valores standardizados DfFit vs standardizados previstos do modelo de Carreiras.

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA**

**Tabela A 10 - Sumário do modelo de regressão de Motoristas
Model Summary^b**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Motoristas	,958 ^a	,917	,916	,01750	1,843

a. Predictors: (Constant), Ev1007, TipStand, dsup5, TipMini, TipMidi, Velcom_Log, ComprMedPerc

b. Dependent Variable: Consumo_Log

**Tabela A 11 - Análise ANOVA do modelo de Motoristas
ANOVA^b**

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1,625	7	,232	758,309	,000 ^a
Motoristas Residual	,147	480	,000		
Total	1,772	487			

a. Predictors: (Constant), Ev1007, TipStand, dsup5, TipMini, TipMidi, Velcom_Log, ComprMedPerc

b. Dependent Variable: Consumo_Log

Tabela A 12 - Coeficientes de regressão do modelo de Motoristas

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
Constante	1,965	,024		82,230	,000		
Logaritmo da velocidade comercial	-,145	,020	-,161	-7,360	,000	,360	2,779
Porcentagem de condução de tipologia Mini	-,005	,000	-,845	-55,150	,000	,736	1,359
Porcentagem de condução de tipologia Standard	-,002	,000	-,561	-33,719	,000	,624	1,604
Porcentagem de condução de tipologia Midi	-,002	,000	-,325	-21,222	,000	,735	1,360
Porcentagem percursos com inclinação superior a 5%	,002	,000	,310	14,390	,000	,373	2,685
Comprimento médio de percurso por motorista	-,002	,000	-,132	-4,906	,000	,237	4,219
Porcentagem de ocorrência do evento 1007	,003	,001	,077	5,689	,000	,943	1,060

a. Dependent Variable: Consumo_Log

Tabela A 13 - Diagnóstico de resíduos do modelo de Motoristas

Residuals Statistics ^a					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1,3399	1,8147	1,6569	,05777	488
Std. Predicted Value	-5,487	2,732	,000	1,000	488
Standard Error of Predicted Value	,001	,007	,002	,001	488
Adjusted Predicted Value	1,3365	1,8143	1,6569	,05772	488
Residual	-,04521	,04504	,00000	,01737	488
Std. Residual	-2,583	2,574	,000	,993	488
Stud. Residual	-2,592	2,592	,000	1,003	488
Deleted Residual	-,04551	,04569	-,00001	,01775	488
Stud. Deleted Residual	-2,608	2,608	,000	1,005	488
Mahal. Distance	,466	87,101	6,986	10,420	488
Cook's Distance	,000	,147	,003	,011	488
Centered Leverage Value	,001	,179	,014	,021	488

a. Dependent Variable: Consumo_Log

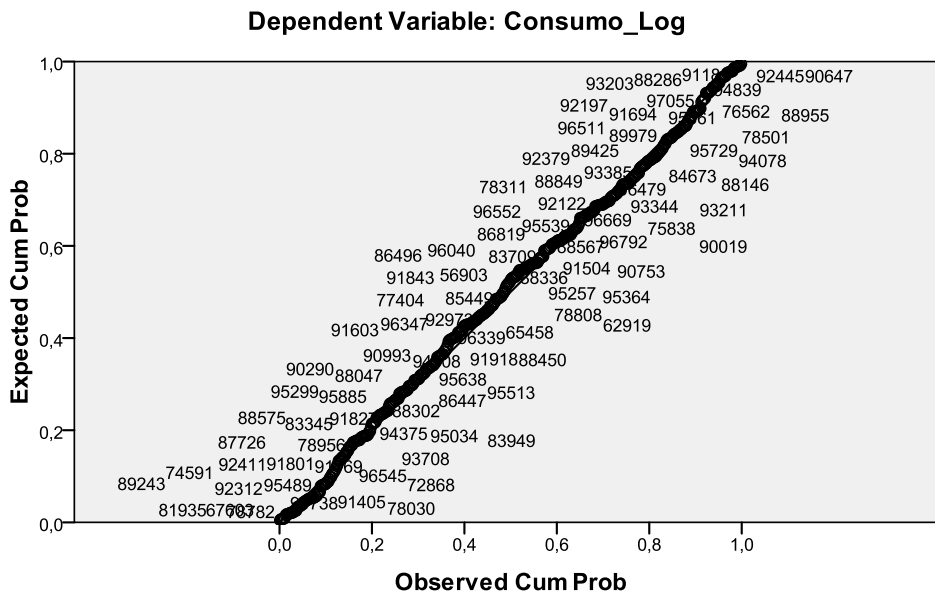


Figura A 6 - Gráfico de probabilidade normal dos resíduos do modelo de Motoristas.

ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA

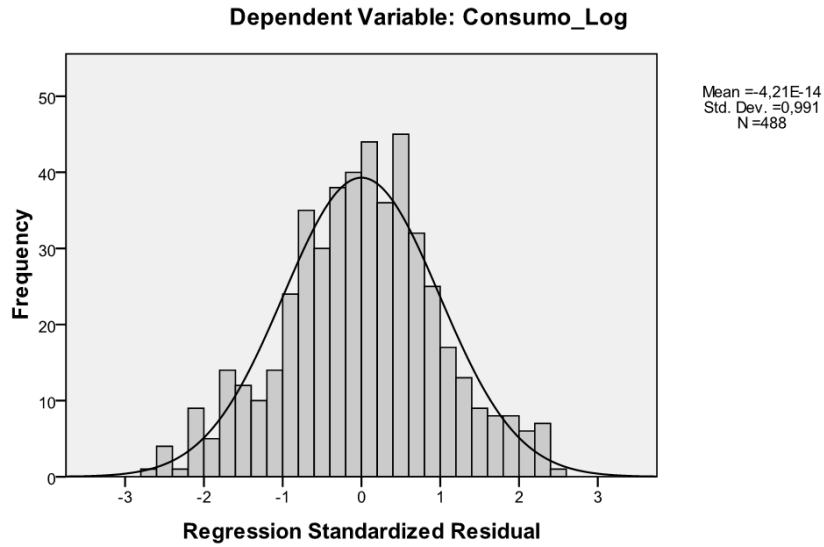


Figura A 7 - Histograma de resíduos do modelo de Motoristas.

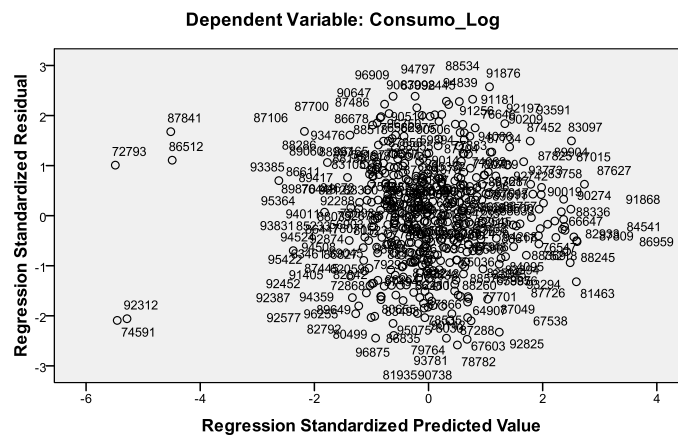


Figura A 8 - Gráfico de resíduos standardizados vs resíduos standardizados previstos do modelo de Motoristas.

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA**

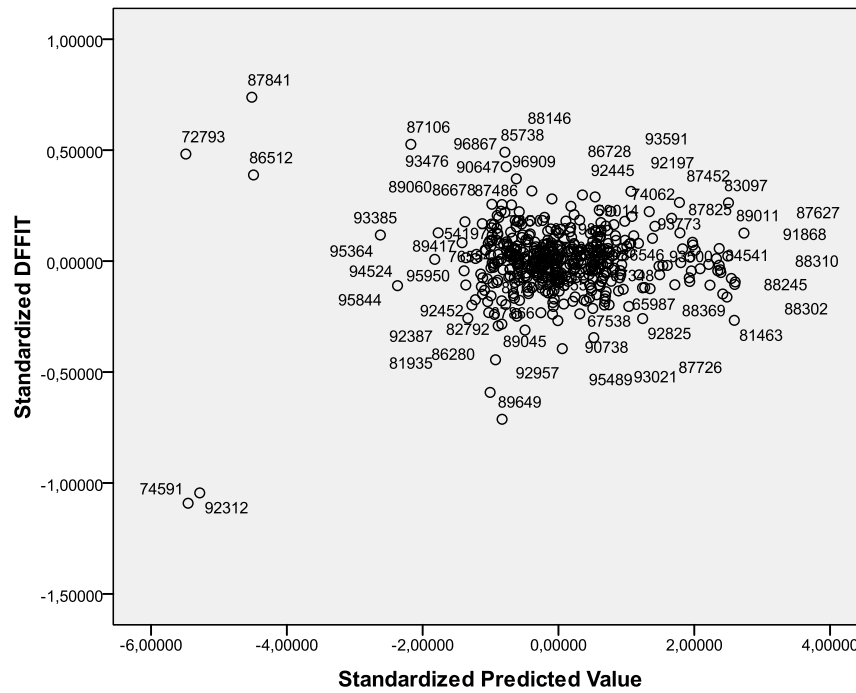


Figura A 10 - Gráfico de valores standardizados Dffit vs standardizados previstos do modelo de Motoristas.

**Tabela A 16 - Sumário do modelo de regressão de Viaturas
Model Summary^b**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Viaturas	,977 ^a	,954	,952	,02393	1,404

a. Predictors: (Constant), IdadeMedMots_Log, Ev1067, VelCom, Massa_Log, Comprmed

b. Dependent Variable: Consumo_Log

Tabela A 17 - Análise ANOVA do modelo de Viaturas

ANOVA^b

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	1,184	5	,237	413,370	,000 ^a
Viaturas Residual	,057	99	,001		
Total	1,240	104			

a. Predictors: (Constant), IdadeMedMots_Log, Ev1067, VelCom, Massa_Log, Comprmed

b. Dependent Variable: Consumo_Log

**Tabela A 18 - Coeficientes de regressão do modelo de Viaturas
Coefficients^a**

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
Constante	-1,737	,141		-12,311	,000		
Logaritmo da massa de cada viatura	,776	,020	,912	38,079	,000	,805	1,242
Velocidade comercial	-,006	,001	-,195	-4,735	,000	,273	3,666
Comprimento médio de percurso	-,003	,001	-,140	-3,325	,001	,261	3,831
Porcentagem de ocorrência do evento 1067	,017	,004	,108	4,838	,000	,931	1,075
Logaritmo da idade média de motoristas	,195	,077	,056	2,516	,013	,927	1,079

a. Dependent Variable: Consumo_Log

Tabela A 19 - Diagnóstico de resíduos do modelo de Viaturas

Residual Statistics					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1,2686	1,8543	1,6586	,10668	105
Std. Predicted Value	-3,656	1,834	,000	1,000	105
Standard Error of Predicted Value	,003	,013	,005	,002	105
Adjusted Predicted Value	1,2633	1,8555	1,6584	,10695	105
Residual	-,05857	,05225	,00000	,02335	105
Std. Residual	-2,448	2,184	,000	,976	105
Stud. Residual	-2,472	2,216	,004	1,001	105
Deleted Residual	-,05975	,05382	,00022	,02460	105
Stud. Deleted Residual	-2,539	2,262	,004	1,009	105
Mahal. Distance	,391	29,561	4,952	5,852	105
Cook's Distance	,000	,096	,009	,014	105
Centered Leverage Value	,004	,284	,048	,056	105

a. Dependent Variable: Consumo_Log

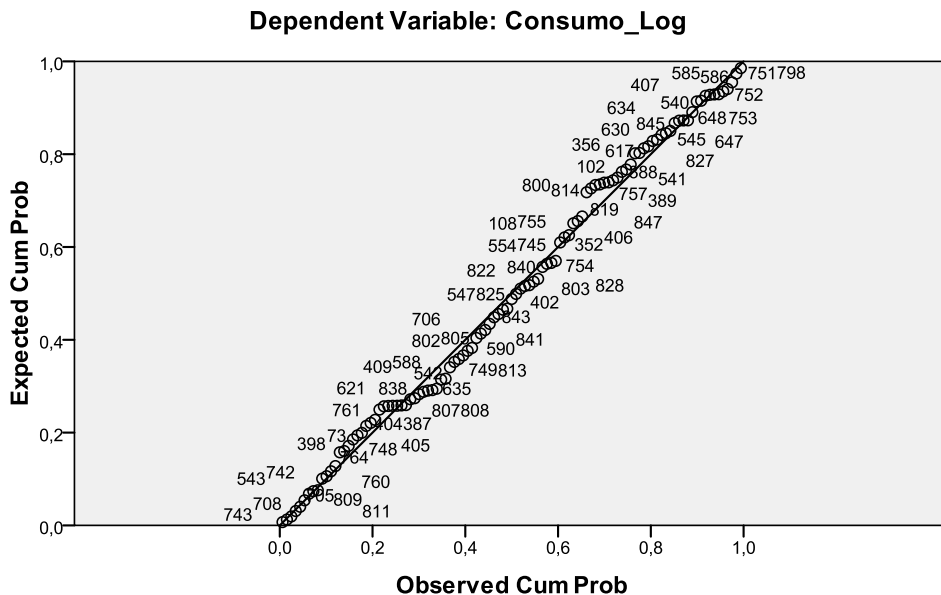


Figura A 11 - Gráfico de probabilidade normal dos resíduos do modelo de Viaturas.

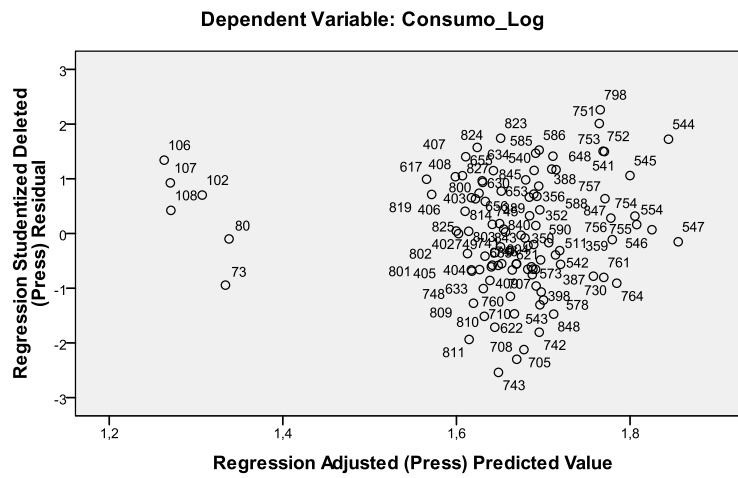


Figura A 14 - Gráfico de resíduos studentizados vs valores ajustados previstos (press) do modelo de Viaturas.

Tabela A 20- Teste de normalidade de resíduos do modelo de Viaturas
One-Sample Kolmogorov-Smirnov Test

		Unstandardized Residual
N		105
Normal Parameters^{a,b}	Mean	,0000000
	Std. Deviation	,02334715
Most Extreme Differences	Absolute	,066
	Positive	,053
	Negative	-,066
Kolmogorov-Smirnov Z		,672
Asymp. Sig. (2-tailed)		,757
Exact Sig. (2-tailed)		,731
Point Probability		,000

a. Test distribution is Normal.

b. Calculated from data.

Tabela A 21 - Teste de normalidade de resíduos do modelo de Viaturas.

Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
Unstandardized Residual	,989	105	,574

**ANEXO III – ELASTICIDADES DO CONSUMO FACE ÀS
VARIÁVEIS SELECCIONADAS NOS MODELOS**

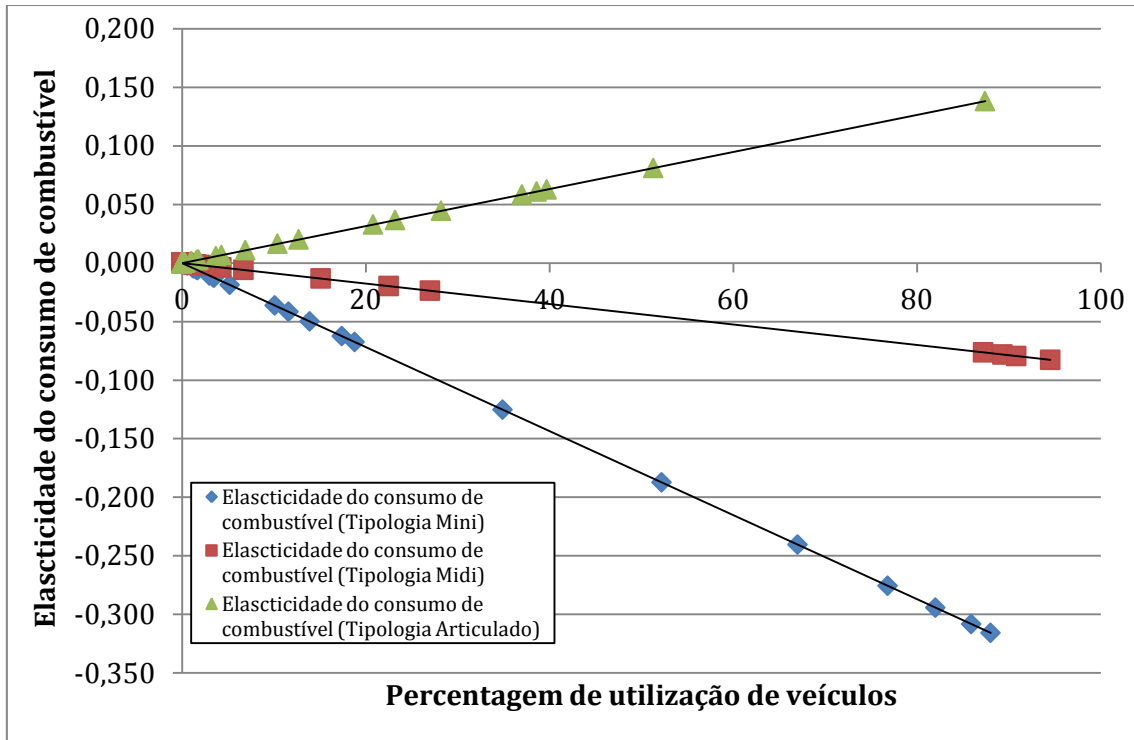


Figura A 16 - Elasticidades do consumo de combustível face às tipologias Mini, Midi e Articulado no modelo de Carreiras.

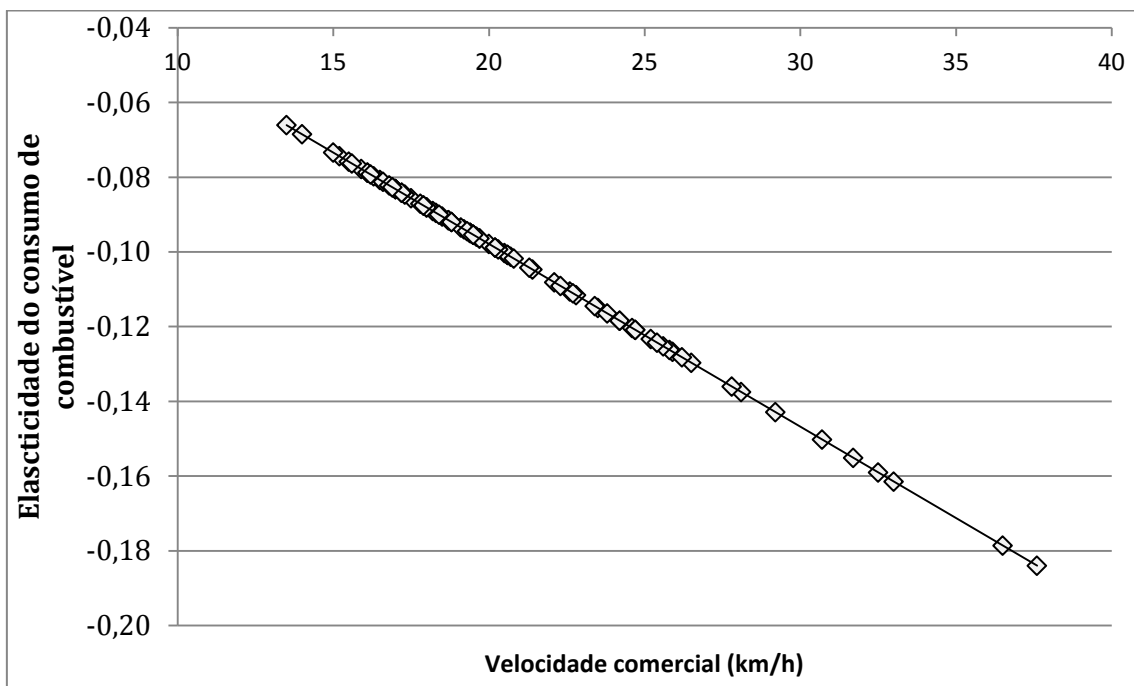


Figura A 17 - Elasticidade do consumo de combustível face à velocidade comercial no modelo de Carreiras.

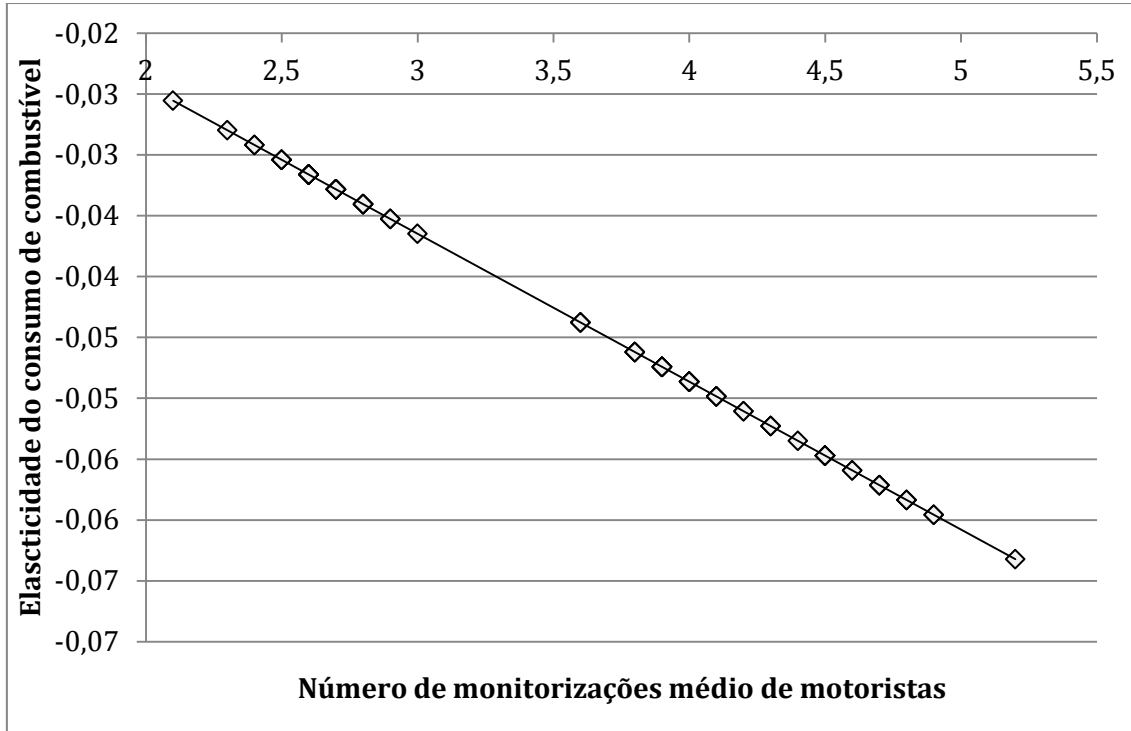


Figura A 18 - Elasticidade do consumo de combustível face ao número de monitorizações médio de motoristas no modelo de Carreiras.

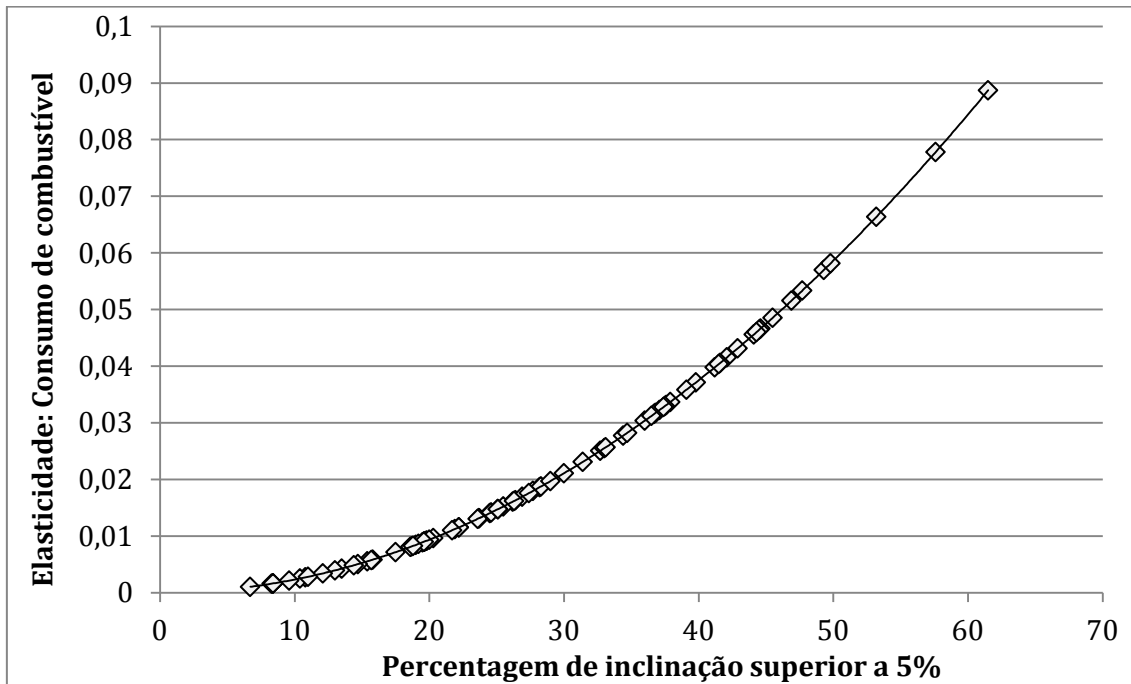


Figura A 19 - Elasticidade do consumo de combustível face à inclinação superior a 5% do modelo de Carreiras.

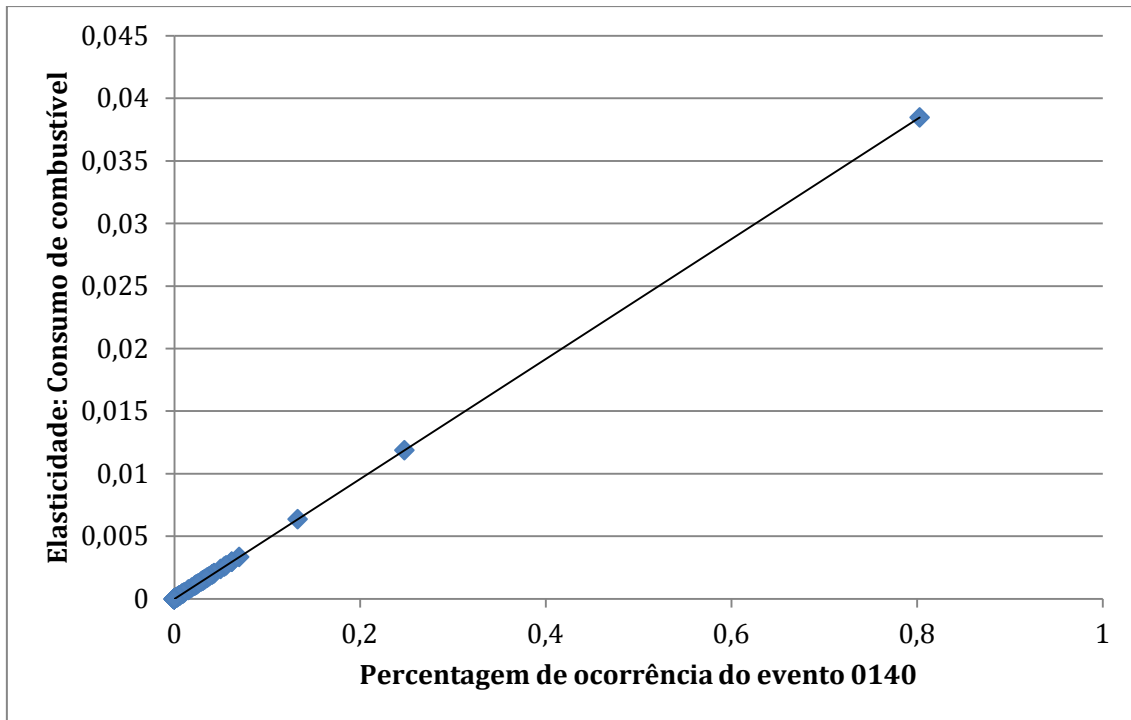


Figura A 20 - Elasticidade do consumo de combustível face à percentagem do evento 0140 no modelo de Carreiras.

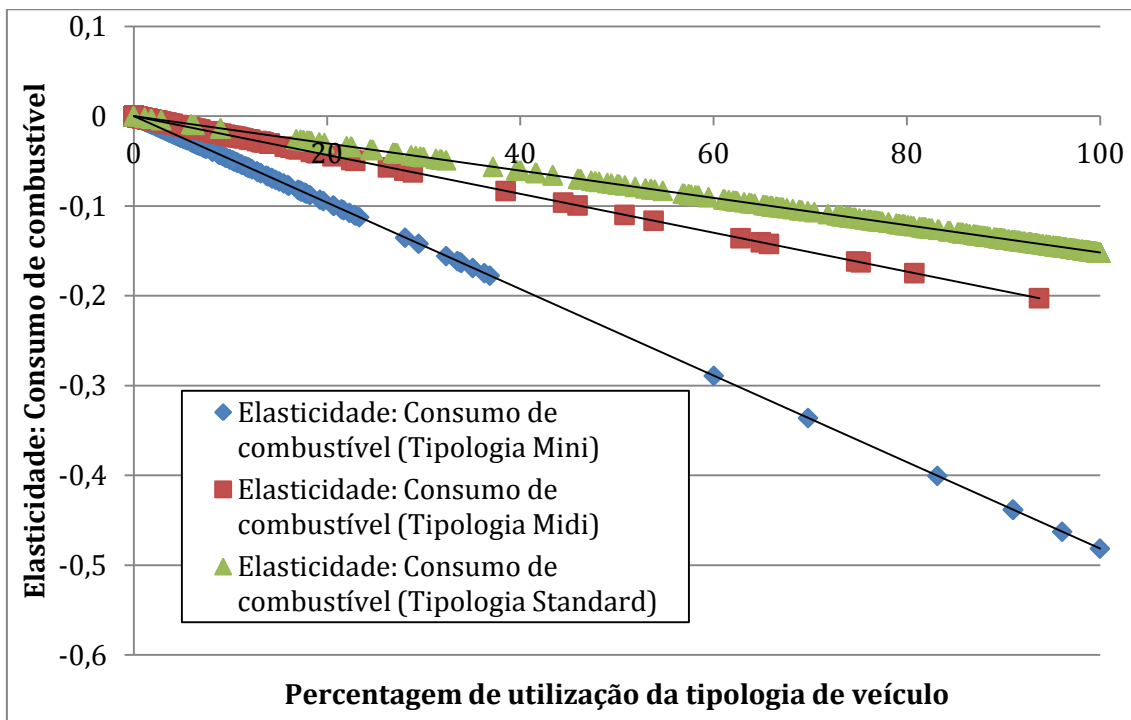


Figura A 21 - Elasticidade das tipologias de viatura Mini, Midi e Standard no modelo de Motoristas.

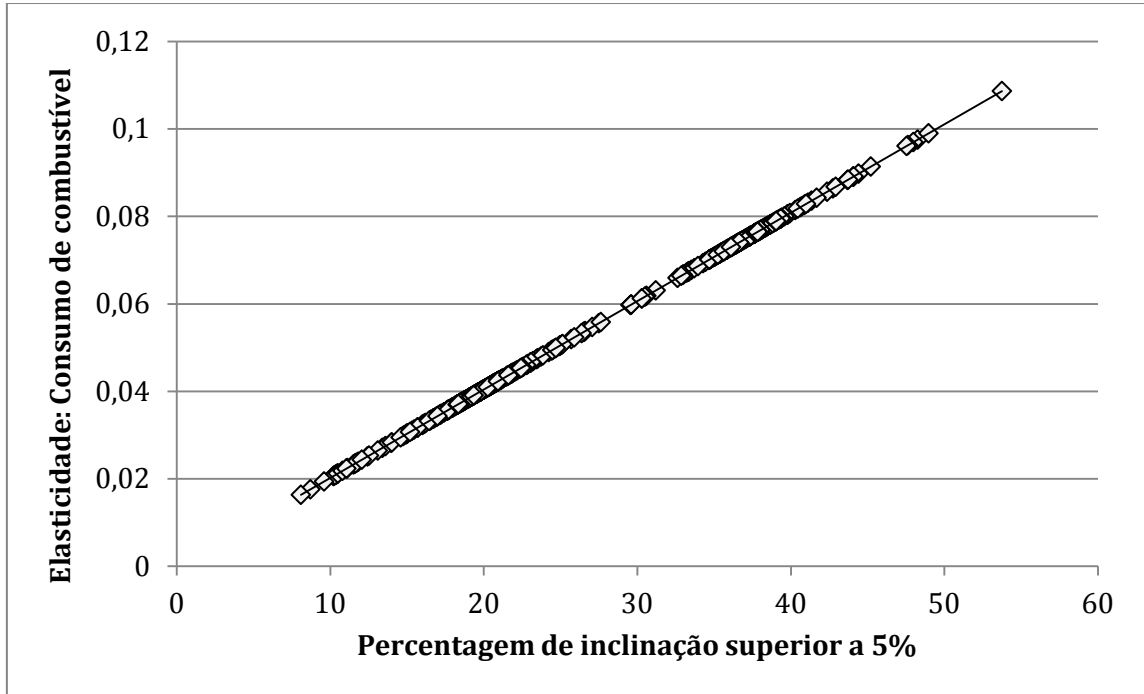


Figura A 22 - Elasticidade do consumo de combustível face à percentagem de inclinação superior a 5% no modelo de Motoristas.

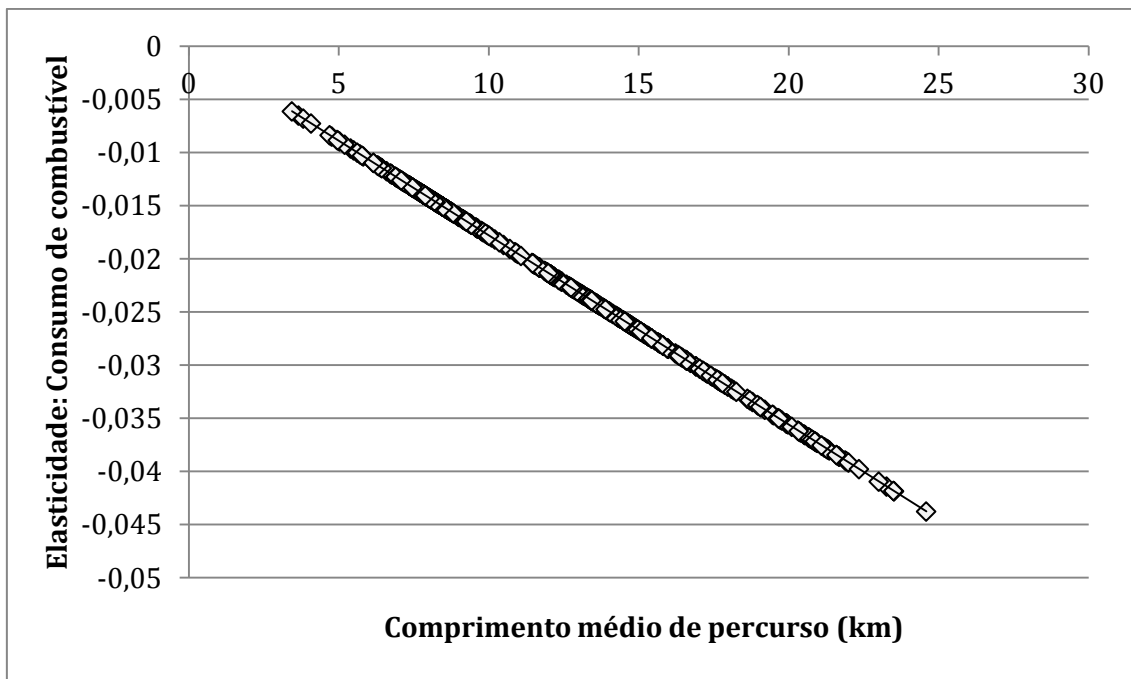


Figura A 23 - Elasticidade do consumo de combustível face ao comprimento médio de percurso no modelo de Motoristas.

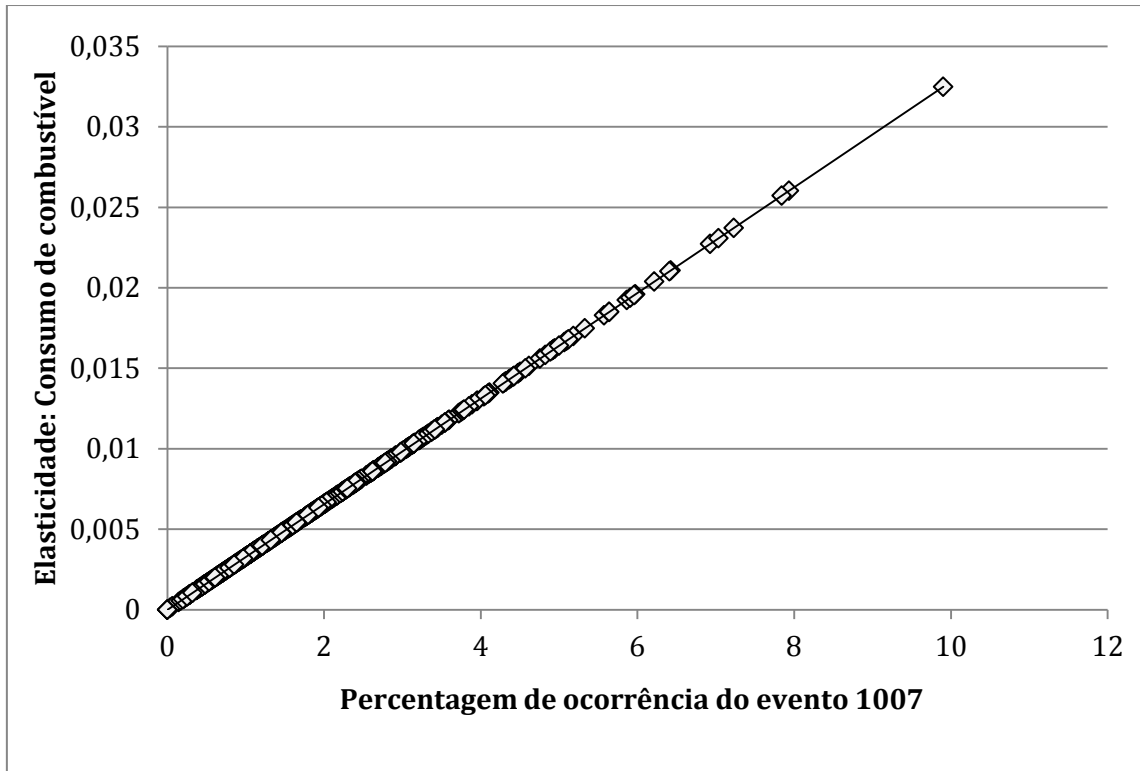


Figura A 24 - Elasticidade do consumo de combustível face à porcentagem de ocorrência do evento 1007 no modelo de Motoristas.

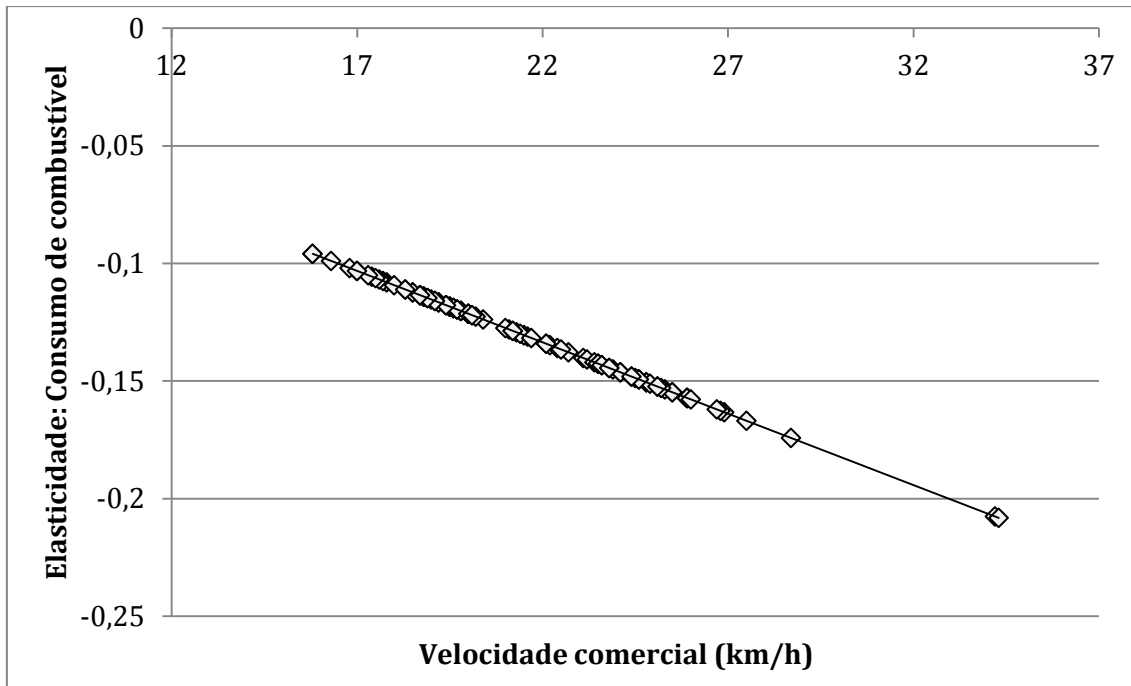


Figura A 25 - Elasticidade do consumo de combustível face à velocidade comercial no modelo de Viaturas.

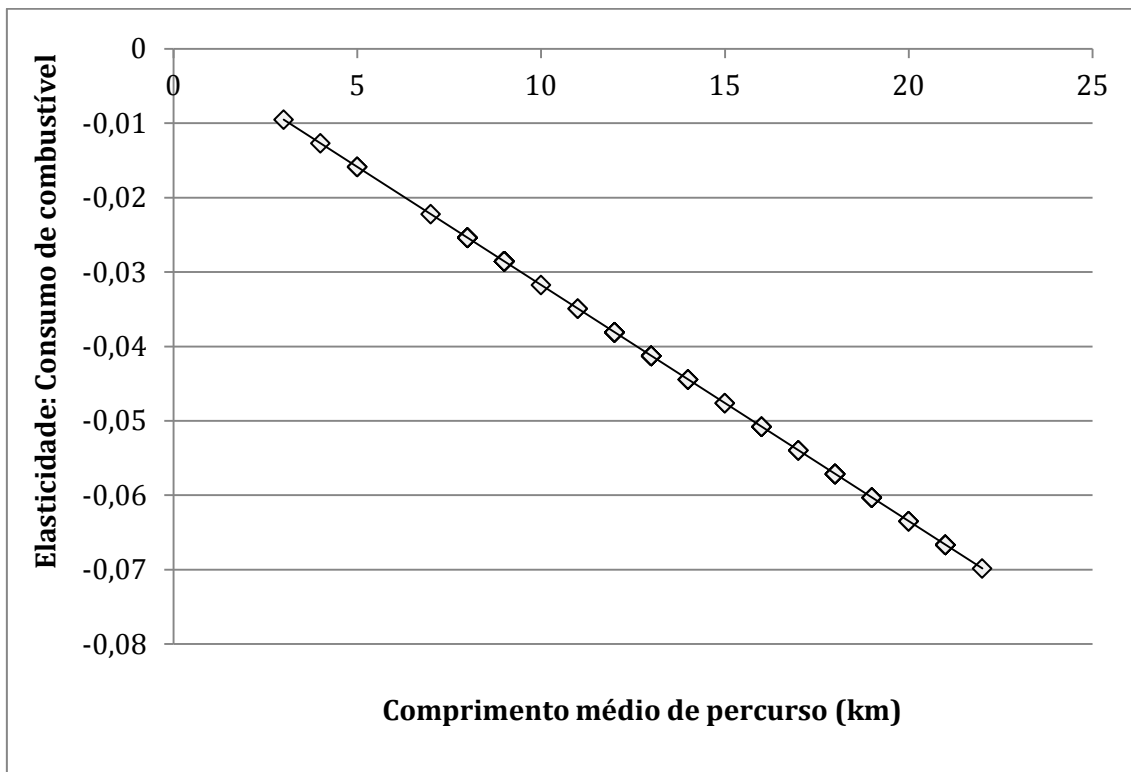


Figura A 26 - Elasticidade do consumo de combustível face ao comprimento médio de percurso no modelo de Viaturas.

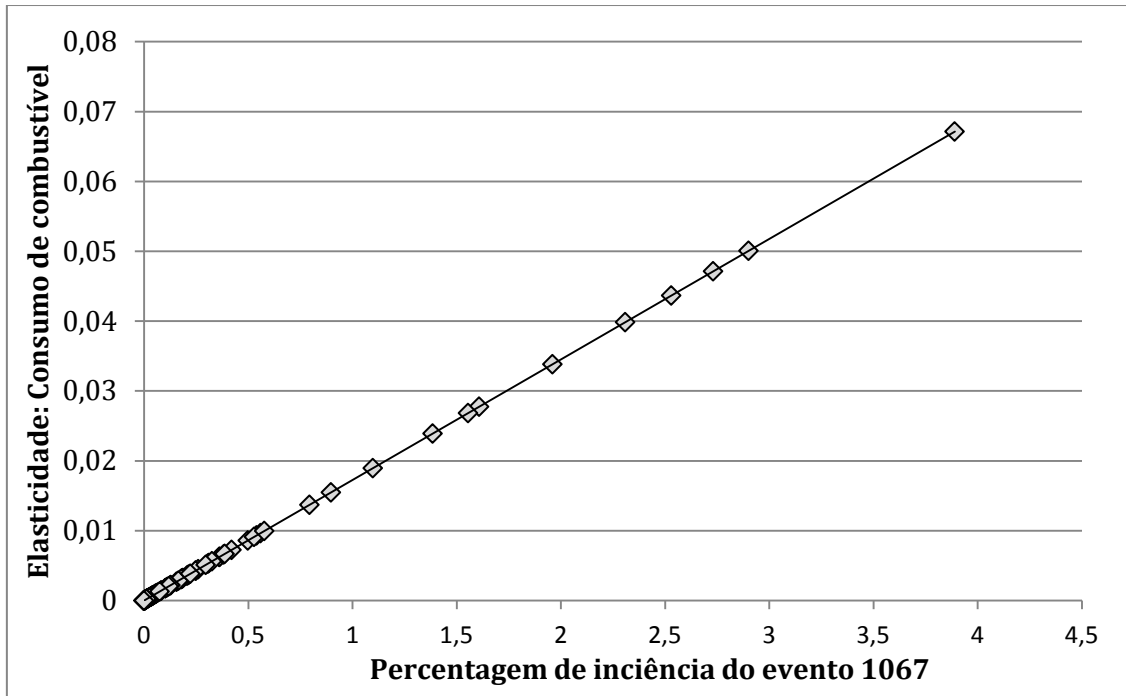


Figura A 27 - Elasticidade do consumo de combustível relativamente à percentagem de ocorrência do evento 1067 no modelo de Viaturas.

**ANEXO IV – CÓDIGO DESENVOLVIDO EM SQL PARA
DESENVOLVIMENTO DE VARIÁVEIS**

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA**

```
select viagens.carr as carreira, viagens.`date` as data, viagens.numviat as veiculo,
avg(consumoscorrig.consumo) as consumo, viagens.kms as kms FROM viagens INNER JOIN consumoscorrig
ON viagens.date = consumoscorrig.date and viagens.numviat = consumoscorrig.numviat group by viagens.carr,
viagens.`date`, viagens.numviat;
```

```
select viagens.carr as carreira, viagens.`date` as data, viagens.numviat as veiculo, avg(consumo2.consumo) as
consumo, viagens.km as kms FROM viagens INNER JOIN consumo2 ON viagens.date = consumo2.date and
viagens.numviat = consumo2.numviat group by viagens.carr, viagens.`date`, viagens.numviat;
```

```
select viagens.carr inner join consumoscorrig ON viagens.date = consumoscorrig.date and viagens.numviat =
consumoscorrig.numviat from viagens, consumoscorrig group by viagens.carr, viagens.`date`, viagens.numviat;
```

```
select id, avg(cons) from consumo group by id;
```

```
select idviat, (sum(consumo*kms))/sum(kms) from cons group by idviat;
```

```
select numviat, (sum(consumo*km))/sum(km) from consumo2 group by numviat into outfile
'VIATconsumo.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.numviat, avg(carreiras.comprim) from viagens, carreiras where viagens.carr = carreiras.idcarr
group by viagens.numviat into outfile 'VIATURAScomprviagens.csv' fields terminated by ';' lines terminated by
'\r\n';
```

```
select viagens.nummot, avg(viagens.km) from viagens group by viagens.nummot into outfile
'MOTORISTAScomprviagens.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.numviat, sum(viagens.km) from viagens group by viagens.numviat into outfile
'VIATkmspercorridos.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select idviat, count(idavar) from avar group by idviat into outfile 'avartotais.csv' fields terminated by ';' lines
terminated by '\r\n';
```

```
select idviat, avg(velcom) from viagens group by idviat;
```

```
select viagens.idviat, sum(carr.comprim) from viagens, carr group by idviat;
```

```
select q1.x as x, q1.y as r1, q2.y as r2 from (select x, avg(y) from w where x % 2 group by x) as q1 inner join
(select x, avg(y) from w where !(x % 2) group by x) as q2 on q1.x = q2.x;
```

```
select viagens.carr, count(if(event = 100, 1, NULL)) / count(*) * 100 as `100`, count(if(event = 110, 1, NULL)) /
count(*) * 100 as `110`, count(if(event = 120, 1, NULL)) / count(*) * 100 as `120`, count(if(event = 121, 1,
NULL)) / count(*) * 100 as `121`, count(if(event = 130, 1, NULL)) / count(*) * 100 as `130`, count(if(event =
140, 1, NULL)) / count(*) * 100 as `140`, count(if(event = 141, 1, NULL)) / count(*) * 100 as `141`,
count(if(event = 1005, 1, NULL)) / count(*) * 100 as `1005`, count(if(event = 1007, 1, NULL)) / count(*) * 100
as `1007`, count(if(event = 1009, 1, NULL)) / count(*) * 100 as `1009`, count(if(event = 1010, 1, NULL)) /
```

```
count(*) * 100 as `1010`, count(if(event = 1014, 1, NULL)) / count(*) * 100 as `1014`, count(if(event = 1015, 1, NULL)) / count(*) * 100 as `1015`, count(if(event = 1026, 1, NULL)) / count(*) * 100 as `1026`, count(if(event = 1058, 1, NULL)) / count(*) * 100 as `1058`, count(if(event = 1060, 1, NULL)) / count(*) * 100 as `1060`, count(if(event = 1062, 1, NULL)) / count(*) * 100 as `1062`, count(if(event = 1067, 1, NULL)) / count(*) * 100 as `1067`, count(if(event = 1069, 1, NULL)) / count(*) * 100 as `1069`, count(if(event = 1070, 1, NULL)) / count(*) * 100 as `1070`, count(*) as `count` from log group by carr into outfile 'log.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select numviat, count(if(event = 100, 1, NULL)) / count(*) * 100 as `100`, count(if(event = 110, 1, NULL)) / count(*) * 100 as `110`, count(if(event = 120, 1, NULL)) / count(*) * 100 as `120`, count(if(event = 121, 1, NULL)) / count(*) * 100 as `121`, count(if(event = 130, 1, NULL)) / count(*) * 100 as `130`, count(if(event = 140, 1, NULL)) / count(*) * 100 as `140`, count(if(event = 141, 1, NULL)) / count(*) * 100 as `141`, count(if(event = 1005, 1, NULL)) / count(*) * 100 as `1005`, count(if(event = 1007, 1, NULL)) / count(*) * 100 as `1007`, count(if(event = 1009, 1, NULL)) / count(*) * 100 as `1009`, count(if(event = 1010, 1, NULL)) / count(*) * 100 as `1010`, count(if(event = 1014, 1, NULL)) / count(*) * 100 as `1014`, count(if(event = 1015, 1, NULL)) / count(*) * 100 as `1015`, count(if(event = 1026, 1, NULL)) / count(*) * 100 as `1026`, count(if(event = 1058, 1, NULL)) / count(*) * 100 as `1058`, count(if(event = 1060, 1, NULL)) / count(*) * 100 as `1060`, count(if(event = 1062, 1, NULL)) / count(*) * 100 as `1062`, count(if(event = 1067, 1, NULL)) / count(*) * 100 as `1067`, count(if(event = 1069, 1, NULL)) / count(*) * 100 as `1069`, count(if(event = 1070, 1, NULL)) / count(*) * 100 as `1070`, count(*) as `count` from log group by numviat into outfile 'VIATURASlog.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select nummot, count(if(event = 100, 1, NULL)) / count(*) * 100 as `100`, count(if(event = 110, 1, NULL)) / count(*) * 100 as `110`, count(if(event = 120, 1, NULL)) / count(*) * 100 as `120`, count(if(event = 121, 1, NULL)) / count(*) * 100 as `121`, count(if(event = 130, 1, NULL)) / count(*) * 100 as `130`, count(if(event = 140, 1, NULL)) / count(*) * 100 as `140`, count(if(event = 141, 1, NULL)) / count(*) * 100 as `141`, count(if(event = 1005, 1, NULL)) / count(*) * 100 as `1005`, count(if(event = 1007, 1, NULL)) / count(*) * 100 as `1007`, count(if(event = 1009, 1, NULL)) / count(*) * 100 as `1009`, count(if(event = 1010, 1, NULL)) / count(*) * 100 as `1010`, count(if(event = 1014, 1, NULL)) / count(*) * 100 as `1014`, count(if(event = 1015, 1, NULL)) / count(*) * 100 as `1015`, count(if(event = 1026, 1, NULL)) / count(*) * 100 as `1026`, count(if(event = 1058, 1, NULL)) / count(*) * 100 as `1058`, count(if(event = 1060, 1, NULL)) / count(*) * 100 as `1060`, count(if(event = 1062, 1, NULL)) / count(*) * 100 as `1062`, count(if(event = 1067, 1, NULL)) / count(*) * 100 as `1067`, count(if(event = 1069, 1, NULL)) / count(*) * 100 as `1069`, count(if(event = 1070, 1, NULL)) / count(*) * 100 as `1070`, count(*) as `count` from log group by nummot into outfile 'MOTORISTASlog.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
SELECT carr, COUNT(IF(tipo REGEXP "mini", 1, NULL)) / COUNT(*) * 100 AS `1`, COUNT(IF(idavar REGEXP "2", 1, NULL)) / COUNT(*) * 100 AS `2`, COUNT(IF(idavar REGEXP "3", 1, NULL)) / COUNT(*) * 100 AS `3`, COUNT(IF(idavar REGEXP "4", 1, NULL)) / COUNT(*) * 100 AS `4`, COUNT(IF(idavar REGEXP "5", 1, NULL)) / COUNT(*) * 100 AS `5`, COUNT(IF(idavar REGEXP "6", 1, NULL)) / COUNT(*) * 100 AS `6`, COUNT(IF(idavar REGEXP "7", 1, NULL)) / COUNT(*) * 100 AS `7`, COUNT(IF(idavar REGEXP "8", 1, NULL)) / COUNT(*) * 100 AS `8` FROM avarias GROUP BY idviat;
```

```
select viagens.carr as carr, sum(case when viaturas.tipo = 'MINI' then viagens.km end) / sum(viagens.km) * 100 as mini, sum(case when viaturas.tipo = 'MIDI' then viagens.km end) / sum(viagens.km) * 100 as midi, sum(case when viaturas.tipo = 'STANDARD' then viagens.km end) / sum(viagens.km) * 100 as stand, sum(case when viaturas.tipo = 'ARTIC' then viagens.km end) / sum(viagens.km) * 100 as artic from viagens inner join viaturas on viagens.numviat = viaturas.numviat group by carr into outfile 'CARRveicetype.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.nummot as Motorista, sum(case when viaturas.tipo = 'MINI' then viagens.km end) / sum(viagens.km) * 100 as Mini, sum(case when viaturas.tipo = 'MIDI' then viagens.km end) / sum(viagens.km) * 100 as Midi, sum(case when viaturas.tipo = 'STANDARD' then viagens.km end) / sum(viagens.km) * 100 as Stand, sum(case when viaturas.tipo = 'ARTIC' then viagens.km end) / sum(viagens.km) * 100 as Artic from
```

**ANÁLISE AOS VECTORES MAIS INFLUENTES NA PREVISÃO DO CONSUMO DE COMBUSTÍVEL DE UMA EMPRESA
RODOVIÁRIA DE TRANSPORTE DE PASSAGEIROS: A CARREIRA, O MOTORISTA E A VIATURA**

```
viagens inner join viaturas on viagens.numviat = viaturas.numviat group by nummot into outfile 'MOTORISTASveictype.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.numviat as viatura, sum(case when carreiras.tipo = 'URB' then viagens.km end) / sum(viagens.km) * 100 as Urbana, sum(case when carreiras.tipo = 'SUB' then viagens.km end) / sum(viagens.km) * 100 as Suburbana, sum(case when carreiras.tipo = 'DIR' then viagens.km end) / sum(viagens.km) * 100 as Directa from viagens INNER JOIN carreiras on viagens.carr=carreiras.carr group by numviat into outfile 'carreiratype.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.nummot as Motorista, sum(case when carreiras.tipo = 'URB' then viagens.km end) / sum(viagens.km) * 100 as Urbana, sum(case when carreiras.tipo = 'SUB' then viagens.km end) / sum(viagens.km) * 100 as Suburbana, sum(case when carreiras.tipo = 'DIR' then viagens.km end) / sum(viagens.km) * 100 as Directa from viagens INNER JOIN carreiras on viagens.carr=carreiras.idcarr group by nummot into outfile 'MOTORISTAScarreiratype.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.carr as carr, avg(viaturas.vage) as idade, avg(viaturas.massa) as massa from viagens inner join viaturas on viagens.numviat = viaturas.numviat group by viagens.carr into outfile 'vage.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.nummot as motorista, avg(viaturas.massa) as massaviatura from viagens inner join viaturas on viagens.numviat = viaturas.numviat group by viagens.nummot into outfile 'MOTORISTASmassaviaturas.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select carreira, sum(consumo) / sum(km) as consumo from (select viagens.carr as carreira, viagens.`date` as data, viagens.numviat as veiculo, avg(consumo2.consumo) * sum(viagens.km) as consumo, sum(viagens.km) as km, count(viagens.carr and viagens.`date`) as `count` from viagens inner join consumo2 on viagens.date = consumo2.date and viagens.numviat = consumo2.numviat group by viagens.carr, viagens.`date`, viagens.numviat) as q1 group by carreira into outfile 'CARRconsumption.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select motorista, sum(consumo) / sum(km) as consumo from (select viagens.nummot as motorista, viagens.`date` as data, viagens.numviat as veiculo, avg(consumo2.consumo) * sum(viagens.km) as consumo, sum(viagens.km) as km, count(viagens.carr and viagens.`date`) as `count` from viagens inner join consumo2 on viagens.date = consumo2.date and viagens.numviat = consumo2.numviat group by viagens.nummot, viagens.`date`, viagens.numviat) as q1 group by motorista into outfile 'MOTORISTAconsumption.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select consumos2.numviat, avg(consumoscorrig.consumo) as Consumo from consumoscorrig group by numviat into outfile 'VIATURASconsumption.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.carr, avg(delta) / 60 as dm, sum(viagens.km) / sum(delta / 60 / 60) as vcm from viagens group by viagens.carr into outfile 'CARRvcm.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.numviat, avg(delta) / 60 as dm, sum(viagens.km) / sum(delta / 60 / 60) as vcm from viagens group by viagens.numviat into outfile 'VIATvcm.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.nummot, avg(delta) / 60 as dm, sum(viagens.km) / sum(delta / 60 / 60) as vcm from viagens group by viagens.nummot into outfile 'MOTORISTASvcm.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.nummot, sum(km) from viagens group by nummot into outfile 'MOTORISTAS_kms.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.numviat as viatura, avg(carreiras.d0) as d0, avg(carreiras.dinf5) as dinf5, avg(carreiras.dsup5) as dsup5, avg(carreiras.dsup8) as dsup8, avg(carreiras.numparag) as NumParMed, avg(carreiras.distmed) as DistMedPar from viagens inner join carreiras on viagens.carr=carreiras.idcarr group by viagens.numviat into outfile 'declives&carr.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.nummot as viatura, avg(carreiras.d0) as d0, avg(carreiras.dinf5) as dinf5, avg(carreiras.dsup5) as dsup5, avg(carreiras.dsup8) as dsup8, avg(carreiras.numparag) as NumParMed, avg(carreiras.distmed) as DistMedPar from viagens inner join carreiras on viagens.carr=carreiras.idcarr group by viagens.nummot into outfile 'MOTORISTASdeclives&dists.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.carr as carr, avg(motoristas.monit) as monit, avg(motoristas.idade) as idade from viagens inner join motoristas on viagens.nummot = motoristas.nmot group by viagens.carr into outfile 'CARRmonits.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select viagens.numviat as viatura, avg(motoristas.monit) as NumMonit, avg(motoristas.anosexp) as AnosTrab, avg(motoristas.idade) as Idade from viagens inner join motoristas on viagens.nummot=motoristas.nmot group by viagens.numviat into outfile 'outmotoristas.csv' fields terminated by ';' lines terminated by '\r\n';
```

```
select numviat, IF(tipologia REGEXP "min", 1, NULL) , IF(tipologia REGEXP "mid", 1, NULL), IF(tipologia REGEXP "art", 1, NULL) from viatdum group by numviat into outfile 'viaturasdummie.csv' fields terminated by ';' lines terminated by '\r\n';
```