

ESTUDO DA DISTRAÇÃO E SONOLÊNCIA EM MOTORISTAS PROFISSIONAIS COM BASE EM DADOS DE UM SISTEMA DE MONITORIZAÇÃO DO CONDUTOR

Sara Ferreira
Zafeiris Kokkinogenis
António Couto
Universidade do Porto
Faculdade de Engenharia

RESUMO

O objetivo do presente estudo é explorar os dados gerados por um sistema de monitorização do condutor que emite alertas quando deteta eventos de sonolência e distração. Com base na data e hora do início e fim das viagens foi possível caracterizar as viagens tendo em conta a duração das mesmas bem como o número e duração das paragens. Por outro lado, os alertas foram agregados às referidas viagens, considerando-se assim a cada observação (sub-viagem cujo acumulado corresponde ao total da viagem) uma frequência de alertas. Os dados foram obtidos para 70 condutores de seis empresas distintas. Com a aplicação de dois modelos lineares generalizados com distribuição binomial negativa, obteve-se as elasticidades das variáveis que representam assim o efeito das características da viagem na ocorrência de alertas de distração e sonolência. O estudo demonstra o potencial de utilização de dados deste tipo para o estudo do comportamento dos motoristas profissionais.

ABSTRACT

The aim of the present study is to explore data gathered by a driver monitoring system which emits alerts when detect drowsiness and distraction. The journey is characterized based on its duration, number and duration of breaks using the date and hour recorded by the system. The alerts are aggregated by sub-journey (the accumulate of the sub-journeys is a journey), which corresponds to an observation with a specific number of alerts. The data comprehends 70 drivers from six distinct companies. Generalized linear models with negative binomial distribution were applied and elasticities estimated considering two distinct models depending on the type of alert - drowsiness and distraction. Based on these models, the effects of the journey on this type of driver behaviour were obtained. Overall, the study clearly shows the potential of this type of data to analyse de behaviour of professional drivers.

1. INTRODUÇÃO

Comportamentos do condutor do tipo distração e sonolência contribuem para a ocorrência de cerca de 5% a 25% do total dos acidentes (European Commission 2015) e aproximadamente 20% dos acidentes graves e fatais (Connor, et al. 2002; Kecklund, et al. 2011). Estes comportamentos do condutor estão associados a uma diminuição do desempenho na condução bem como das capacidades cognitivas (Atchley, et al. 2017; Fitzharris, et al. 2017; Sweeney, et al. 1995). No caso dos motoristas profissionais estes comportamentos têm repercussões mais graves na medida em que os mesmos conduzem durante longos períodos de tempo e como rotina diária estando como tal mais expostos ao risco de ocorrência de acidentes. De facto, os autores (Chen, et al. 2016) reportaram que a fadiga e sonolência na condução são uma das principais causas de acidentes envolvendo veículos pesados. Já a distração está associada a cerca de 70% dos acidentes do mesmo tipo segundo vários autores (Chen, et al. 2016; European Commission 2015; Fitzharris, et al. 2017; Lemke, et al. 2016; Sparrow, et al. 2016). Consequentemente, estes comportamentos do condutor têm vindo a ser analisados por diversos autores utilizando diferentes metodologias de investigação quer em ambiente

simulado (simulador de condução) quer em ambiente real (estudos naturalísticos).

É o caso por exemplo do estudo naturalístico realizado pelos autores Hanowski, et al. (2008) com a participação de 103 condutores cujo objetivo foi avaliar o desempenho da condução através da quantificação de eventos críticos de segurança (ECS). Na mesma linha Soccolich, et al. (2013) implementaram um estudo naturalístico envolvendo 97 condutores para avaliar o desempenho na condução. Os dois estudos analisaram o número de ECS nas diferentes horas de condução até um máximo de 11 horas (limite máximo legal) tendo os autores observado não existir diferenças significativas nas últimas horas de condução. No estudo de Soccolich, et al. (2013) foi ainda analisado o impacto das paragens na condução ao longo da viagem tendo os autores concluído que parar pode reduzir a taxa de ECS no período a seguir à paragem, independentemente se essa paragem foi para atividades profissionais ou de descanso. Já os estudos naturalísticos realizados por Sparrow, et al. (2016) e por Chen, et al. (2016) focaram-se na avaliação do sono e o seu impacto no desempenho da condução nos motoristas profissionais. Embora com abordagens distintas, os dois estudos concluíram da importância do sono para mitigar a fadiga e consequente melhor desempenho na condução.

Os estudos acima descritos tiveram como foco a sonolência e fadiga, no entanto, também a distração dos condutores profissionais foi analisada considerando o mesmo tipo de abordagem, ou seja, estudos naturalísticos. É o caso do estudo de Hanowski, et al. (2005) em que foram observados 41 condutores de veículos pesados em ambiente real, tendo sido analisados os ECS. Dos 2737 eventos registados, 178 foram atribuídos à distração do condutor. Estes últimos eventos foram analisados tendo sido identificados 34 tipos de distração distintos.

Para estudar o caso específico da distração pelo uso do telemóvel, Hickman, et al. (2014) analisaram ECS registados durante um estudo naturalístico com 207 condutores profissionais de veículos pesados comerciais e autocarros. Este estudo foi a continuação de um estudo prévio de Hickman e Hanowski (2012) cujos resultados apontaram que o uso de telemóvel durante a condução não aumentou a probabilidade de estar envolvido num ECS, sugerindo portanto que esta tarefa não pode ser analisada de forma dicotómica (sim/não). De facto, o estudo de Hickman, et al. (2014) demonstrou que o tipo de utilização do telemóvel influencia o risco de ECS. Por exemplo, falar/ouvir ao telemóvel durante a condução não foi associado, em geral, a um significativo envolvimento em ECS (em alguns casos até diminuiu), enquanto que tarefas do tipo digitar o número, escrever mensagens, etc. foram associadas a um aumento significativo no envolvimento de ECS. Considerando o mesmo tipo de abordagem de investigação, Olson, et al. (2014) analisaram o impacto da distração dos condutores de veículos comerciais utilizando dados de ECS de dois prévios estudos naturalísticos. Os resultados indicaram que 71% dos acidentes, 46% de quase-acidentes e 60% de todos os ECS estavam associados a tarefas não relacionadas com a condução. Os autores concluíram ainda que tarefas complexas ou tarefas que impliquem o desvio do olhar para a estrada durante muito tempo estão associadas a um aumento significativo do risco.

Em alternativa aos estudos naturalísticos foram desenvolvidos estudos com utilização de um simulador de condução. Foi o caso do estudo dos autores Stavrinou, et al. (2016) que teve como objetivo analisar o efeito de ler/escrever mensagens e emails no telemóvel, considerando ainda a autoavaliação dos condutores, no desempenho da condução de veículos pesados comerciais. Os resultados do estudo revelaram que as tarefas que requerem maior

atenção visual estavam associadas a um menor desempenho da condução e que, em geral, à um enviesamento na autoavaliação dos condutores sendo estes mais otimistas em relação à realidade. Também com base num estudo em ambiente virtual, Sollins, et al. (2014) analisaram o impacto da utilização do telemóvel, MP3 *touch* e distrações ambientais externas no desempenho da condução de veículos pesados comerciais. Os resultados indicaram que o desempenho da condução diminuiu em conjunto com um aumento da alocação da atividade cognitiva às distrações, e em particular às resultantes da utilização do telemóvel e MP3 *touch*. Com foco na sonolência, Hallvig, et al. (2013) concluíram que os resultados obtidos em simulador são comparáveis aos obtidos em ambiente real. Contudo, os resultados absolutos indicam que os níveis de sonolência são maiores nos ambientes simulados do que em ambiente real. Como tal, deve-se ter em atenção este facto na generalização dos resultados obtidos em simulador para ambiente real.

Como se pode depreender pelos estudos acima descritos, resultados relevantes foram obtidos com base em estudos naturalísticos e em ambiente simulado. Contudo, a implementação de estudos com estas abordagens é bastante dispendiosa em tempo e custo monetário e como tal nem sempre possível de se considerar. Por outro lado, com o recente e rápido desenvolvimento de novas tecnologias aplicadas à condução e monitorização do condutor, às quais estão associadas a recolha e armazenamento de dados, surge a oportunidade de explorar estes mesmos dados oferecendo uma abordagem complementar às já existentes.

No caso da distração e da sonolência têm vindo a ser desenvolvidos sistemas de monitorização do condutor (SMC) tendo sido publicado diversos estudos com foco na análise da precisão do sistema e/ou no desenvolvimento de metodologias para melhorar o desempenho desses mesmos sistemas (Dumitrescu, et al. 2016; Jackson, et al. 2016; Kawanaka, et al. 2013; Masala and Grosso 2014). Contudo, aparentemente poucos estudos têm sido realizados com base nos dados gerados e armazenados por estes SMC tal como parece indicar o estudo de Atchley, et al. (2017) sobre distração na condução que, tendo como base a análise de 50 anos de investigação sobre o tema, não reporta nenhum estudo que utilize os dados referidos.

É no entanto de salientar o estudo recente de Fitzharris, et al. (2017) que utilizou dados de um SMC implementado em ambiente real em condutores profissionais de veículos pesados com o objetivo de analisar quer a incidência e duração da fadiga quer o impacto do feedback do sistema de monitorização e/ou do operador da frota. O estudo foi realizado em colaboração com uma empresa transportadora e com base em dados históricos dos eventos de fadiga extraídos da base de dados do fornecedor do SMC. Estes dados incluem a data e hora, localização e velocidade dos veículos no instante do evento, tipo de evento, ID do veículo e do operador, tempo em movimento e tempo parado bem como data e hora do início/fim da viagem. Para o estudo, os autores definiram 3 períodos distintos: 1) período de referência em que o SMC regista os dados, mas não emite alerta ao condutor ou empresa; 2) período de alerta em que o SMC emite alerta sonoro ao condutor; 3) período de alerta e feedback em que para além da emissão do alerta, a empresa através de um operador comunica com o condutor dando feedback ao mesmo. O objetivo de definir estes três períodos foi o de avaliar a redução nos eventos de fadiga resultante do SMC e do feedback do operador. Os resultados do estudo na verdade mostraram que a percentagem de eventos de fadiga numa viagem diminuiu de 3,74% durante o primeiro período para 1,3% e 0,18% nos períodos de alerta e de alerta e feedback, respetivamente. Os autores concluíram assim que o estudo evidenciou a relevância

da monitorização dos condutores em tempo-real e como tal deve ser considerada como medida em políticas de segurança.

Tal como no estudo de Fitzharris, et al. (2017), o presente estudo baseia-se em dados históricos de alertas emitidos por um SMC desenvolvido e implementado por uma empresa comercial do setor tecnológico. A empresa facultou os dados armazenados de acordo com a estrutura de registo do SMC que foi implementado em diversos veículos de diferentes empresas (transportadoras, comerciais e operacionais). O SMC deteta eventos de sonolência e distração, emitindo um alerta sonoro ao condutor. O SMC regista dados do evento (por exemplo, data e hora) bem como da viagem (data e hora do início e fim). Assim, é possível associar a cada viagem o número de eventos registados. Cada viagem tem características distintas tais como tempo total de viagem, número de paragens, duração da paragem, etc. Por outro lado, as viagens são realizadas por diferentes condutores em veículos pesados ou ligeiros pertencentes a seis empresas distintas. Considerando os alertas como eventos críticos de segurança (ECS), estes são agregados por viagem e analisados com base em modelos linear generalizados. Desta forma, é possível analisar e identificar fatores de risco associados a estes eventos, separando a distração e sonolência em dois modelos distintos.

Apresenta-se na secção seguinte a descrição dos dados utilizados para a realização do estudo com base na metodologia descrita na secção 3. A secção 4 mostra os resultados dos dois modelos desenvolvidos, sendo a discussão dos mesmos apresentada na secção 5. Finalmente, apresenta-se as conclusões finais com a sugestão de futuros desenvolvimentos.

2. DESCRIÇÃO DOS DADOS

Os dados considerados neste estudo foram fornecidos pela empresa que desenvolveu o SMC tendo sido tratados para garantir a confidencialidade dos condutores. De notar que os autores do presente estudo não supervisionaram a implementação do SMC nem mesmo a estruturação e gestão dos dados. A seguir descreve-se separadamente a fonte de dados e respetivo tratamento bem como a análise estatística das variáveis utilizadas.

2.1. Fonte e tratamento dos dados

A empresa comercial que forneceu os dados para o presente estudo opera no setor dos transportes desenvolvendo tecnologia biométrica de reconhecimento facial. O SMC foi desenvolvido para detetar estados de sonolência e distração, emitindo um alerta sonoro para avisar o condutor. A solução tecnológica desenvolvida compreende uma câmara de infravermelhos de suporte fixo que obtém imagens faciais do condutor. A câmara de vídeo tem uma resolução 480x752 (WxH) e capacidade de gerar 15 imagens por segundo (fps). De acordo com a descrição fornecida pela empresa, o sistema utiliza técnicas avançadas de *deep learning* para análise do olhar e da face. O sistema classifica os alertas como distração ou sonolência sem distinção do tipo de distração (por exemplo, escrever mensagens, olhar para exterior, etc.) nem do tipo de sonolência (por exemplo, episódio de microsonolência, bocejar). A empresa realizou diversos testes gravando imagens vídeos de vários participantes a conduzir em diferentes horas do dia e em diferentes condições viárias. As imagens de vídeo foram analisadas e os eventos validados. A empresa reporta para os algoritmos de visão computacional uma precisão (taxa de verdadeiros positivos) de 92,13% e *recall* (taxa de verdadeiros negativos) de 88,53%.

O sistema regista o início e fim da viagem através da data, hora e local (coordenadas GPS).

Durante a viagem, regista a data, hora e local (GPS) do alerta, o tipo de alerta (distração ou sonolência) e a velocidade instantânea do veículo. Para o presente estudo foi realizado um exaustivo tratamento dos dados que incluiu eliminar dados incompletos ou em duplicado e analisar inconsistências e eventuais erros. Posteriormente, os alertas foram agregados por viagem tendo em conta a informação da data/hora dos alertas e das viagens para cada condutor. A informação dos alertas relativa à localização e velocidade instantânea não foi utilizada neste estudo.

A base de dados corresponde a 70 veículos/condutores de seis empresas distintas que operam em diferentes setores tais como transporte de mercadorias, serviços comerciais e operação de gestão de infraestrutura rodoviária. Os dados foram registados durante a atividade profissional de rotina dos condutores nos horários de trabalho habitual. Tendo em conta que os condutores trabalham em atividades profissionais muito distintas, é de esperar que se observe diferentes tipos de viagens, isto é, com duração e número de paragens distintas, por exemplo. Contudo, os autores não tiveram acesso a qualquer tipo de informação relativa ao planeamento e horário das atividades profissionais realizadas durante a recolha da base de dados.

2.2. Descrição estatística dos dados

A base de dados utilizada neste estudo corresponde ao período entre dezembro de 2015 e setembro de 2016 e compreende 1560 observações de 70 condutores. Estas observações são viagens ou sub-viagens (se o condutor parar durante a viagem, esta é dividida em sub-viagens e como tal, o número de sub-viagens é igual ao número de paragens menos 1) para as quais é determinado o número de alertas. Para cada sub-viagem (por simplificação designa-se daqui para a frente sub-viagem a cada observação) corresponde um ID do condutor e ID da empresa bem como data/hora de início e fim da sub-viagem. Para este estudo optou-se por considerar o tempo e não a distância da viagem, na medida em que o tempo, que se relaciona com a distância e a velocidade do veículo, representa melhor, na opinião dos autores, o efeito da viagem no comportamento do condutor. Por outro lado, o tempo de viagem tem sido também utilizado noutros estudos tal como descrito anteriormente, facilitando assim a comparação dos resultados.

A sequência de sub-viagens contíguas, às quais correspondem os mesmos IDs de condutor e de empresa, identifica uma viagem. O tempo entre sub-viagens da mesma viagem é designado de duração da paragem. Contudo, em alguns casos observou-se valores de duração entre sub-viagens muito elevados para ser considerado como uma paragem da mesma viagem. Acresce o facto de se desconhecer os horários das atividades profissionais dos condutores como tal, não é possível aferir se de facto foi uma paragem ou o fim e início de uma nova viagem. Nesse sentido, estabeleceu-se um valor limite acima do qual o tempo entre sub-viagens representa não uma paragem, mas o fim e início de uma nova viagem. Assim, considerou-se o valor limite de 9 horas tendo como base o Regulamento (CE) n.º 561/2006 do Parlamento Europeu e do Conselho, de 15 de março de 2006, relativo à harmonização de determinadas disposições em matéria social no domínio dos transportes rodoviários que estabelece o mínimo de período de repouso diário reduzido de 9 horas. O mesmo documento estabelece que o período de repouso diário normal é de 11 horas. Com a aplicação do valor limite de 9 horas, as viagens analisadas ficam assim caracterizadas de forma mais estandardizada. De referir que das 1560 observações, 377 apresentavam durações entre sub-viagens superior a 9 horas.

Tendo em conta os dois tipos de eventos analisados, separou-se a análise comportamental em distração e sonolência tendo-se para tal criado duas bases de dados distintas, mas com as mesmas variáveis. As variáveis incluídas são: 1) tempo de duração da sub-viagem (em horas); 2) tempo de duração das paragens (em horas); tempo de duração da viagem (acumulado das sub-viagens); 4) número de paragens realizadas durante a viagem (acumulado de cada sub-viagem); 5) número de alertas por sub-viagem (variável dependente). De notar que na primeira sub-viagem de cada viagem (ou na viagem sem nenhuma paragem, ou seja, com uma sub-viagem) o número de paragens é zero e conseqüentemente também a duração da paragem. A correlação entre as variáveis foi analisada, tendo-se identificado o valor mais alto de 0,7 entre as variáveis número de paragens e tempo da viagem (acumulado das sub-viagens). De facto, é de esperar que quanto maior o tempo da viagem, maior será o número de paragens. Tendo presente esta correlação, foi criada uma nova variável dividindo o número de paragens pelo tempo da viagem, e designando-a de frequência de paragens na viagem. Por fim, considerou-se ainda variáveis dummies a identificar as 6 empresas verificando-se a seguinte distribuição de observações por empresa: empresa 1 (E1) 34%; empresa 2 (E2) 6%; empresa 3 (E3) 29%; empresa 4 (E4) 16%; empresa 5 (E5) 9% e empresa 6 (E6) 6%. As empresas 1 e 6 operam no setor de transportes de mercadoria; as empresas 2, 4 e 5 desenvolvem atividades comerciais de vários tipos; a empresa 3 opera como gestora de infraestrutura rodoviária.

As duas bases de dados incluem 1560 sub-viagens (observações) que correspondem ao total de 447 viagens. A Tabela 1 descreve sumariamente a descrição estatística das variáveis utilizadas. O total do número de alertas de sonolência e distração são de 10374 e 12766, respetivamente. De referir que cerca de 70% e 63% das sub-viagens não registaram nenhum alerta de sonolência e distração, respetivamente. Por outro lado, 10% das viagens foram realizadas sem paragens, ou seja, só com uma sub-viagem. Da Tabela 1 destaca-se que cada sub-viagem tem uma duração em média de 0,8 horas para um total de 1261 horas de condução. Por outro lado, uma viagem demora em média 3,4 horas consistindo em 3,5 sub-viagens em média.

Tabela 1: Descrição estatística das variáveis

Variável	Média	Desvio Padrão	Mínimo	Máximo	Percentil 75	Percentil 90	Percentil 95
Nº alertas de distração por sub-viagem	8,2	37,9	0	514	1,0	9,0	32,0
Nº alertas de sonolência por sub-viagem	6,7	30,9	0	505	0	8,0	27,1
Duração da sub-viagem [h]	0,8	1,0	0,02	7,2	0,7	2,3	3,5
Duração da viagem [h]	3,4	4,9	0,02	35,0	3,5	9,0	13,4
Duração da paragem [h]	0,8	1,6	0	8,9	0,6	2,4	4,6
Nº de paragens por sub-viagem	3,3	5,2	0	39	3,0	8,0	12,0
Frequência de paragens (nº de paragens por sub-viagem)	1,4	2,1	0	20,8	1,4	3,6	4,6

Da Tabela 1 pode-se verificar que, pese embora o número elevado de alertas dos dois tipos, estes concentram-se num número reduzido de observações, tal como sugere os percentis calculados para estas variáveis. É de salientar que o valor máximo de duração de uma sub-viagem (viagem em condução contínua) é de 7,2 horas e, portanto, bastante superior ao valor máximo regulamentado pelo Regulamento (CE) n. o 561/2006 cujo valor é de 4,5 horas. Na verdade, 1% das sub-viagens têm duração superior a 4,5 horas. O mesmo regulamento define que o tempo diário de condução não deve exceder 9 horas, podendo ocasionalmente ser excedido (não mais de duas vezes por semana) até um máximo de 10 horas. Contudo, a duração máxima observada por viagem foi de 35 horas (ver Tabela 1) sendo que 4% das viagens têm duração superior a 10 horas. Cerca de 90% das viagens foram realizadas com pelo menos uma paragem sendo que em média realizaram 3,3 paragens. Ao dividir o número de paragens pelo tempo de condução (i.e. sub-viagem), obtém-se 1,4 paragens por hora. A Tabela 1 indica ainda que em média o condutor pára durante 48 minutos para descansar ou realizar atividades profissionais.

A Tabela 2 descreve as mesmas variáveis para cada uma das seis empresas, de forma a caracterizar as viagens e a ocorrência de alertas tendo em conta a atividade profissional desenvolvida em cada empresa.

Tabela 2: Descrição estatística das variáveis por empresa

	Nº de sub-viagens	Média duração da viagem [h]	Média do nº de paragens por viagem	Média da duração da paragem [h]	Média do nº de alertas de distração por sub-viagem	Média do nº de alertas de sonolência por sub-viagem
Empresa 1	531	1,3	5,9	0,6	2,4	14,6
Empresa 2	93	1,0	1,8	0,5	4,3	5,0
Empresa 3	447	0,8	1,9	1,2	23,4	3,8
Empresa 4	249	0,6	1,5	0,8	2,2	1,5
Empresa 5	140	0,4	4,0	0,7	0,0	0,1
Empresa 6	100	1,1	1,2	0,7	0,6	0,9

As empresas 1 e 6 que operam no setor dos transportes de mercadorias apresentam os valores mais elevados da média da duração da viagem – 1,3 e 1,1 horas, respetivamente. Contudo, as mesmas empresas apresentam valores muito distintos entre elas do número médio de paragens por viagem, ou seja, 5,9 e 1,2 paragens, respetivamente, embora com durações próximas - 36 e 42 minutos, respetivamente. Por outro lado, a empresa 5 destaca-se por ter o valor mais baixo da duração da viagem (24 minutos em média) sendo a segunda empresa com o valor mais elevado de paragens por viagem (4 paragens em média). Estes valores podem denotar o tipo de atividade profissional do condutor que neste caso é comercial. De facto, os valores apresentados na Tabela 2 deverão estar associados ao tipo de atividade profissional do condutor de cada empresa, refletindo assim o propósito e os horários associados à condução destes profissionais. Finalmente, é de salientar a diferença observada nos valores médios da frequência dos alertas dos dois tipos.

3. METODOLOGIA

À semelhança do estudo dos acidentes em que habitualmente se considera a frequência dos acidentes como variável dependente e são aplicados modelos lineares generalizados (MLG) com distribuição Poisson ou binomial negativa (BN), também a análise da frequência dos alertas agregada por sub-viagem será realizada com base nos mesmos modelos. De facto, o

número de alertas por sub-viagem corresponde a um valor inteiro positivo sendo que se verifica a existência de um número elevado de zeros (70% e 63% das observações de alertas de sonolência e distração, respetivamente).

Assim, segundo a metodologia habitual da análise de frequência de acidentes, testou-se a hipótese da variância ser igual à média aplicando testes de sobredispersão para as duas bases de dados separadamente, tendo-se concluído que a distribuição BN representa melhor os dados em estudo. A função densidade de probabilidade da distribuição BN pode ser representada com a seguinte formulação (Washington, et al. 2003):

$$P(Y = y_i | \mu_i, \alpha) = \frac{\Gamma(y_i + (1/\alpha))}{\Gamma(1/\alpha) \cdot y_i!} \left(\frac{1/\alpha}{(1/\alpha) + \mu_i} \right)^{1/\alpha} \left(\frac{\mu_i}{(1/\alpha) + \mu_i} \right)^{y_i} \quad (1)$$

em que α : Parâmetro de sobredispersão a ser estimado na regressão BN;
 $\Gamma(\cdot)$: Função gama;
 μ_i : Valor médio $\mu_i = E\{Y\}$.

Por outro lado, o valor médio relaciona-se com as variáveis explicativas através de uma função de ligação logarítmica:

$$\mu_i = \exp(\beta X_i) \text{ ou equivalente } \ln(\mu_i) = \beta X_i \quad (2)$$

em que β : Parâmetros a serem estimados;
 X_i : Vetor de variáveis explicativas.

Os parâmetros estimados β permitem inferir acerca do efeito das variáveis explicativas na ocorrência dos alertas. Para analisar os valores estimados, as elasticidades podem ser calculadas da seguinte forma (Washington, et al. 2003):

$$E_{x_{ik}}^{\mu_i} = \frac{\partial \mu_i}{\mu_i} \times \frac{x_{ik}}{\partial x_{ik}} = \beta_k x_{ik} \quad (3)$$

em que E : Elasticidade;
 x_{ik} : Valor de variável explicativa k^{th} para a observação i ;
 β_k : Valor estimado para a variável explicativa k^{th} ;
 μ_i : Valor esperado da frequência para a observação i .

Esta elasticidade pode ser interpretada como o efeito na frequência esperada μ_i de 10% de aumento do valor da variável explicativa.

Para aferir sobre qual a melhor forma função representativa da relação entre cada variável explicativa e a variável dependente, considerou-se a aplicação do método *Integrate-Differentiate* proposto por Hauer e Bamfo (1997). Simplificadamente, o método consiste em desenhar a função integral empírica para cada variável e comparar com a forma gráfica de funções conhecidas do tipo gama, exponencial, potência etc. A função com a forma mais semelhante é selecionada sendo posteriormente analisada a respetiva função acumulada. Se esta corresponder a uma linha então pode-se concluir que se ajusta aos dados analisados.

Através da aplicação do referido método concluiu-se que no caso da base de dados de sonolência, a função potência representa a relação entre a variável dependente e as variáveis duração da sub-viagem, duração da viagem, e número de paragens. Para a duração da paragem a função exponencial foi selecionada. Assim, o modelo BN com a função de ligação logarítmica pode ser representada da seguinte forma:

$$\ln \mu_i = \beta_0 + \beta_1 \ln(DurSubV_i) + \beta_2 \ln(DurViag_i) + \beta_3 \ln(FreqPar_i) + \beta_4 DurPar_i + \beta_{5-9} Empresa_i \quad (4)$$

em que μ_i : Média do número de alertas de sonolência da observação i ;
 β_k : Parâmetros estimados ($k=0$ a 9).

No caso da base de dados de alertas de distração, o mesmo método sugere que as relações entre as variáveis explicativas e a variável dependente podem ser representadas da seguinte forma:

$$\ln \mu_i = \beta_0 + \beta_1 \ln(DurSubV_i) + \beta_2 \ln(DurViag_i) + \beta_3 \ln(FreqPar_i) + \beta_4 \ln(DurPar_i) + \beta_{5-9} Empresa_i \quad (5)$$

em que μ_i : Média do número de alertas de distração da observação i ;
 β_k : Parâmetros estimados ($k=0$ a 9).

De notar que o objetivo deste estudo não é o de prever o número de alertas, mas antes analisar o efeito das variáveis na frequência de alertas.

4. RESULTADOS

Os resultados dos dois modelos são apresentados a seguir nas Tabelas 3 e 4. A Tabela 3 apresenta os resultados do modelo 1 dos alertas de sonolência e a Tabela 4 apresenta os resultados do modelo 2 dos alertas de distração, considerando as formas funcional das equações (4) e (5), respetivamente. A empresa E1 foi considerada a categoria de referência para as variáveis dummies das empresas.

Os primeiros resultados mostraram que a variável dummy da empresa E3 no modelo 1, bem como as variáveis duração da viagem e duração da paragem no modelo 2, não são estatisticamente significantes. Optou-se por eliminar estas mesmas variáveis nos resultados finais de cada modelo, sendo que as restantes variáveis são estatisticamente significantes com 95% de confiança, tal como se pode observar pelas Tabelas 3 e 4.

Da Tabela 3 pode-se constatar que os parâmetros estimados para as categorias E4, E5 e E6 apresentam valores estimados distintos, mas todos com sinal negativo, ao contrário da categoria E2 cujo sinal é positivo para o valor estimado. O valor positivo estimado para o parâmetro da variável duração da sub-viagem indica que aumentando o tempo de condução contínua, aumenta o número de alertas de sonolência. Por outras palavras, um aumento de 10% do tempo de duração da sub-viagem corresponde a um aumento de 15% dos alertas, *ceteris paribus*. Da mesma forma se pode concluir que um aumento de 10% da duração da viagem ou do número de paragens por hora de viagem corresponde a um aumento de 2% ou 9% dos alertas, respetivamente. Contudo, se o número de paragens se mantiver fixo, a elasticidade da duração da viagem passa a ser positiva de valor +0.767, significando que um aumento de 10% da duração da viagem aumenta em 7,7% os alertas. Por último, o sinal negativo do parâmetro estimado para a duração da paragem revela que um aumento do tempo

de paragem decresce o número de alertas. Contudo, sendo esta variável representada pela função exponencial, o efeito da elasticidade não é constante pois depende do valor considerado para a variável ($E_{x_{ik}}^{\mu_i} = -0.087(DurPar_i)$).

Tabela 3: Resultados do modelo BN – alertas de sonolência

Parâmetro	Coefficiente	Desvio Padrão	P[Z>z]
β_0	1,650	0,111	0,0000
β_1	1,518	0,093	0,0000
β_2	-0,156	0,042	0,0000
β_3	-0,923	0,158	0,0000
β_4	-0,087	0,027	0,0011
$\beta_5 - E_2$	0,913	0,235	0,0001
$\beta_6 - E_4$	-0,597	0,196	0,0024
$\beta_7 - E_5$	-1,841	0,347	0,0000
$\beta_8 - E_6$	-1,059	0,205	0,0000
α	5,129	0,321	0,0000

1560 observações; log likelihood = -2210,4372; Mc
Fadden Pseudo R² = 0,8447

Os resultados da Tabela 4 mostram que os parâmetros estimados para as categorias que representam as empresas são estatisticamente significantes e com valores e sinais distintos entre si. Como tal, os alertas de distração ocorrem de forma distinta em função da empresa do condutor. O valor positivo estimado para a duração da sub-viagem indica que a um aumento de 10% do tempo de condução contínua corresponde um aumento de 10% do número de alertas de distração. Pelo contrário, o valor negativo do parâmetro estimado para o número de paragens por hora indica que os alertas decrescem 3% com o aumento de 10% desta mesma variável.

Tabela 4: Resultados do modelo BN – alertas de distração

Parâmetro	Coefficiente	Desvio Padrão	P[Z>z]
β_0	0,276	0,141	0,0496
β_1	0,970	0,072	0,0000
β_3	-0,295	0,129	0,0222
$\beta_5 - E_2$	1,301	0,312	0,0000
$\beta_6 - E_3$	3,485	0,211	0,0000
$\beta_7 - E_4$	1,299	0,204	0,0000
$\beta_8 - E_5$	-2,921	0,529	0,0000
$\beta_9 - E_6$	-0,720	0,250	0,0040
α	4,646	0,273	0,0000

1560 observações; log likelihood = -2626,1998; Mc
Fadden Pseudo R² = 0,7731

5. DISCUSSÃO E CONCLUSÕES FINAIS

Dos resultados obtidos pode-se concluir que o aumento do tempo de condução contínua aumenta o risco de ocorrência de alertas quer distração quer de sonolência. Mais especificamente, um aumento de 10% desse tempo, aumenta em 15% e 10% os alertas de sonolência e distração, respetivamente. Na verdade, esta variável pode ser interpretada como

variável de exposição à ocorrência de alertas tal como a variável do volume de tráfego o é para os acidentes. Estes resultados estão na linha do estudo de Soccolich, et al. (2013) que demonstrou que existe uma relação positiva entre o tempo de condução e a ocorrência de ECS. Assim, quer o presente estudo quer estudos anteriores demonstram a importância de definir um limite para o tempo de condução contínua.

É interessante de observar que se se considerar o número de paragens por hora de viagem constante, o tempo de duração da viagem (acumulado de 3,5 sub-viagens em média) tem um efeito negativo na ocorrência de alertas de sonolência. Contudo, é de referir que o aumento do tempo de viagem deverá estar associado a um aumento do número de paragens tal como indicado pelo elevado valor de correlação entre as mesmas, tornando constante a frequência de paragens. Por outro lado, e no caso do modelo dos alertas de distração, a variável duração da viagem não é estatisticamente significativa tal como a variável duração da paragem. Este resultado poderá sugerir que no caso de comportamentos de distração como sejam ler mensagens, desvira olhar da estrada, etc. podem não ser afetados pelo tempo total da viagem (acumulado das sub-viagens mais a duração das paragens) desde que o condutor páre ocasionalmente e independentemente do tempo que demore estas paragens.

Já os eventos de sonolência, associados à fadiga do condutor, são expectáveis de diminuir se o condutor interromper a viagem para descansar, comer ou exercer atividades profissionais. Os resultados do modelo apontam de facto para esse efeito pois aumentando o número de paragens por hora de viagem, o número de alertas diminui cerca de 9%. Mais uma vez estes resultados estão em linha com os do estudo de Soccolich, et al. (2013) que demonstrou que após a paragem da condução, verificou-se uma redução de 28-50% da taxa de ECS. Para além deste efeito da paragem, o presente estudo mostra que o aumento da duração dessa mesma paragem reduz o número de alertas de sonolência proporcionalmente ao valor dessa mesma duração ($-0.087 \times DurParag$). O mesmo estudo de Soccolich, et al. (2013) não analisou o efeito da duração da paragem embora os autores concluíssem que interromper a condução para atividades de qualquer tipo (descansar ou trabalhar) diminui a taxa de ocorrência de ECS. Já o efeito negativo da duração da viagem (acumulado de sub-viagens) poderá captar o efeito distinto das várias sub-viagens refletindo uma diminuição do efeito ao longo do acumulado, eventualmente devido à influência de outras variáveis não incluídas no modelo tais como, hora do dia, tipo de via, volume de tráfego, etc.

Assim, e no caso dos eventos de sonolência, pode-se concluir que o aumento da duração da condução contínua aumenta o número de eventos de sonolência, mas o aumento do tempo da viagem (condução descontínua) diminui os eventos de sonolência. Esta conclusão deve ser interpretada com cuidado na medida em que contradiz o regulamento europeu que define um limite de tempo para a viagem diária. Contudo, é ainda de referir nesta matéria que os estudos de Soccolich, et al. (2013) e de Hanowski, et al. (2008) apontam para resultados semelhantes ao presente estudo na medida em que não observaram diferenças no risco de envolvimento em ECS nas últimas horas da condução em relação às primeiras. Os autores concluem que o estudo não suporta a hipótese de que o risco aumenta durante a condução de um veículo comercial na décima primeira hora em relação às horas prévias de condução. A justificação que os autores sugerem é que a maior exposição ao tráfego/conflitos, a complexidade de condução e inércia poderá ser maior nas primeiras horas.

O presente estudo abrangeu a análise de condutores de diferentes empresas. Das dummies incluídas nos modelos a identificar cada uma das seis empresas, cinco demonstraram ter efeitos distintos na ocorrência de alertas. Este resultado sugere que estas variáveis poderão capturar o efeito do tipo de viagem associada a cada empresa (horários, tipo de trajeto, propósito da viagem, etc.). Na verdade, as empresas analisadas estão associadas, em geral, a atividades de negócio distintas. Por outro lado, estas mesmas variáveis apresentaram resultados distintos dependendo do tipo de evento, isto é, distração ou sonolência, salientando que estes comportamentos têm origens distintas.

Os resultados obtidos estão, em geral, em linha com os resultados obtidos em estudos prévios, acrescentando, no entanto, uma nova dimensão aos efeitos da viagem no comportamento do condutor. Apesar das evidentes limitações do estudo que resultam fundamentalmente por ter uma abordagem “oportunista”, isto é, sem qualquer controlo do ambiente real monitorizado, o estudo demonstrou o potencial de explorar este tipo de dados. Estudos deste tipo poderão demonstrar e identificar diferentes padrões de viagens, condutores ou empresas e como tal orientar o desenvolvimento de tecnologias ajustadas a esses padrões.

REFERÊNCIAS BIBLIOGRÁFICAS

- Atchley, Paul, Ashleigh V. Tran, and Mohammad Ali Salehinejad
2017 Constructing a publically available distracted driving database and research tool. *Accident Analysis & Prevention* 99, Part A:306-311.
- Chen, Guang Xiang, et al.
2016 The influence of daily sleep patterns of commercial truck drivers on driving performance. *Accident Analysis & Prevention* 91:55-63.
- Connor, Jennie, et al.
2002 Driver sleepiness and risk of serious injury to car occupants: population based case control study. *Bmj* 324(7346):1125.
- Dumitrescu, C, et al.
2016 Developing a multi sensors system to detect sleepiness to drivers from transport systems. *Design and Technology in Electronic Packaging (SIITME)*, 2016 IEEE 22nd International Symposium for, 2016, pp. 175-178. IEEE.
- European Commission
2015 Driver Distraction. European Commission, Directorate General for Transport
- Fitzharris, M., et al.
2017 The relative importance of real-time in-cab and external feedback in managing fatigue in real-world commercial transport operations. *Traffic Injury Prevention* 18:S71-S78.
- Hallvig, David, et al.
2013 Sleepy driving on the real road and in the simulator—A comparison. *Accident Analysis & Prevention* 50:44-50.
- Hanowski, R.J., et al.
2008 Analysis of risk as a function of driving-hour: assessment of driving hours 1 through 11. Federal Motor Carrier and Safety Administration, USDOT.
- Hanowski, Richard J., Miguel A. Perez, and Thomas A. Dingus
2005 Driver distraction in long-haul truck drivers. *Transportation Research Part F: Traffic Psychology and Behaviour* 8(6):441-458.
- Hauer, Ezra, and Joseph Bamfo
1997 Two tools for finding what function links the dependent variable to the explanatory variables. *ICTCT Conference, Lund, Sweden, 1997. Proceedings of ICTCT 97.*
- Hickman, J. S., and R. J. Hanowski
2012 An Assessment of Commercial Motor Vehicle Driver Distraction Using Naturalistic Driving Data. *Traffic Injury Prevention* 13(6):612-619.
- Hickman, J. S., R. J. Hanowski, and J. Bocanegra
2014 Distraction in commercial trucks and buses: Assessing prevalence and risk in conjunction with

- crashes and near-crashes. *In* Distracted Driving in Commercial Vehicles and Buses: Research and Analyses. Pp. 223-286.
- Jackson, Melinda L., et al.
2016 The utility of automated measures of ocular metrics for detecting driver drowsiness during extended wakefulness. *Accident Analysis & Prevention* 87:127-133.
- Kawanaka, Haruki, et al.
2013 Identification of Cognitive Distraction Using Physiological Features for Adaptive Driving Safety Supporting System. *International Journal of Vehicular Technology* 2013:18.
- Kecklund, Göran, et al.
2011 Sleepiness and the risk of car crash: a case control study.
- Lemke, Michael K., et al.
2016 Understanding the role of sleep quality and sleep duration in commercial driving safety. *Accident Analysis & Prevention* 97:79-86.
- Masala, G. L., and E. Grosso
2014 Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses. *Transportation Research Part C: Emerging Technologies* 49:32-42.
- Olson, R. L., et al.
2014 Driver distraction in commercial vehicle operations. *In* Distracted Driving in Commercial Vehicles and Buses: Research and Analyses. Pp. 1-222.
- Socolich, Susan A., et al.
2013 An analysis of driving and working hour on commercial motor vehicle driver safety using naturalistic data collection. *Accident Analysis & Prevention* 58:249-258.
- Sollins, B., et al.
2014 Truck driving distractions: Impact on performance and physiological response. *Proceedings of the Human Factors and Ergonomics Society*, 2014. Vol. 2014-January, pp. 2171-2175.
- Sparrow, Amy R., et al.
2016 Naturalistic field study of the restart break in US commercial motor vehicle drivers: Truck driving, sleep, and fatigue. *Accident Analysis & Prevention* 93:55-64.
- Stavrinos, D., et al.
2016 Commercial Truck Driver Health and Safety. *Workplace Health and Safety* 64(8):369-376.
- Sweeney, Margaret M., et al.
1995 Need for sleep: discriminating between fatigue-related and nonfatigue-related truck accidents. *Proceedings of the Human Factors and Ergonomics Society*, 1995. Vol. 2, pp. 1122-1126.
- Washington, Simon, Matthew Karlaftis, and Fred Mannering
2003 *Statistical and econometric methods for transportation data analysis*. Boca Raton London New York Washington, D.C.: Chapman & Hall/CRC.