- Article type: paper
- Date text written or revised: **30 January 2016, revised 27 February 2017, final 20 March 2017**
- 5139 words, 4 figures

A tool to predict environmental risk to GB rail infrastructure

Author 1

- Jason Sadler, BSc
- Principal Research Fellow, GeoData, University of Southampton, Southampton, UK

Author 2

- Oleksandr Kit, MSc, PhD
- Senior Research Fellow, GeoData, University of Southampton, Southampton, UK

Author 3

- Jeremy Austin, BSc
- Senior Research Fellow, GeoData, University of Southampton, Southampton, UK

Author 4

- David Griffin, BA, MSc, PhD
- Principal System Safety Engineer, RSSB, London, UK

Contact: Jason Sadler, GeoData, University of Southampton, Highfield, Southampton, SO17 1BJ, UK. jds@geodata.soton.ac.uk 44 (0)2380 594628

Abstract

Researchers from the University of Southampton have collaborated with RSSB on the GeoSRM pilot project to analyse and map derailments, suicides and slips/trips/falls risk across the GB rail network, initially in the Wessex region. This area of research has been extended to incorporate the impact of historic and "real time" environmental data on rail risk. A detailed description of these investigations is presented, along with the resulting methodology and a prototype toolkit which incorporates environmental conditions combined with rail incident data to better model and predict increased risk in real time.

Keywords

Environment; Railway systems; Risk & probability analysis

List of notation

A_i	is the rainfall raster for ith day
μ_i	is the coefficient for weighting ith day's rainfall
B	is the long-term average monthly rainfall raster
C	is the soil clay content raster
D	is the rail network mask raster
E	is GeoSRM's derailment risk model raster
λ_j	are balancing coefficients
I_n	Intervals used by styling rules for the output rasters

1. Introduction

Computer-assisted modelling and assessment of risks on Britain's railways are important components of optimisation of maintenance costs and safety assurance strategies. These risks vary geospatially across the rail network due to variations in timetable, asset condition, linespeed, etc. However, within the transportation field, geospatial analysis of risk has largely been confined to analysis of road risk e.g. Chen and Zhou (2016) and Dezman *et al* (2016). This is due to the greater number of incidents occurring on the road network being more amenable to statistical analysis. Within the rail industry, where data permits, geospatial modelling has also been used, e.g. for modelling trespasser injuries in the United States (Wang *et al*, 2016).

Where sufficient historical data does not exist, due to the sparsity of accident data, predictive methods may be used instead. Gheorghe *et al* (2005) developed a risk assessment for dangerous goods that could be used to identify risk "hot spots" and display them through a decision support system (DSS) on a map. Their approach used fault and event trees and master logic diagrams although did not take account of train running (i.e. timetables). Building on these ideas, the University of Southampton (GeoData) has collaborated with RSSB since 2013 on a project to develop a Geospatial Risk Model (GeoSRM). The GeoSRM, described in detail by Sadler *et al.* (2016), incorporates both the predictive fault and event tree modelling for the low frequency events (train derailments) alongside statistical modelling for the high frequency events (passenger slip, trip and falls at stations and suicide by member of the public). It does this through an interactive web mapping interface, based both on the geospatial properties of the network as well as the operational (timetable) information.

With expertise in the GB rail infrastructure and safety risk model, RSSB have developed risk models and offline tools to calculate underlying risk values (Dennis and Somaiya, 2004). GeoData specialise in the management and visualisation of large, complex volumes of geospatial data, using open standards (well-defined and documented data formats which support free, vendor-neutral access to underlying documents or records) and open source software tools (those published under a license which permits free redistribution and modification of source code). The team has been responsible for the online system, including the database architecture and web mapping interface. The project has evolved and developed through a high level of interaction between the two groups, and with other rail industry experts and stakeholders.

The GeoSRM pilot models risk based on the location of assets, timetables, topography, socio-economic data and previous event statistics. It was perceived, however, that a wealth of high resolution environmental data was available, both historic and "real time", which had the potential to improve the modelling of rail risk, particularly for derailments caused by structural problems within the underlying soils and by 'trees on the line' events.

A NERC 'Probability, Uncertainty and Risk in the Environment' (PURE) research grant was awarded to investigate this potential further and to extend the GeoSRM pilot to develop a tool which demonstrates the application of open environmental data to the improvement of risk modelling for the GB rail network.

A detailed description of the work is presented in this paper, beginning in Section 2 with an outline of the background research and justification of the assumptions and data sources used for risk calculations. Section 3 describes the software architecture, technical implementation details and formulae used for risk score calculation. Section 4 deals with the results validation process and discusses advantages and limitations of the approach. Sections 5 and 6 outline directions for future work and present conclusions.

2. Environmental risk components

The study began with a review of relevant existing research, including meetings with a number of rail industry stakeholders, earthworks experts and academic partners and projects, such as geotechnical engineers from the Transport Research Group at the University of Southampton, and members of the iSMART project and FutureNet programme. An extensive survey was undertaken of potential environmental data, assessing usefulness, accessibility, spatial/temporal resolution and extent, as well as problems and limitations presented by their paucity.

The development component of the project set out to integrate high resolution precipitation data, long-term climatic averages and soil clay content data, with location data for tracks, assets and incidents, provided by RSSB and Network Rail. This could be tested for the whole of the UK, while within the Wessex region (the existing GeoSRM pilot) there was the additional availability of modelled derailment risk, which could be combined with the recent environmental data to deliver improved topical risk calculations of derailment risk for the region, with potential for future nationwide extension.

The risk of incidents on railway lines is a function of a broad range of factors, such as line maintenance history, infrastructure resilience or line usage statistics. Apart from man-made factors, the short- and long-term condition of the natural environment in the vicinity of tracks and their interaction in time and space play an important role in the composition of risk levels, as these directly influence stability of soils under tracks and adjacent slopes and are frequently the main underlying cause of derailments and damage to the railway infrastructure.

Earthworks are of special interest, predominantly due to the diversity and structural properties of the substrates and used materials. The greatest risk to the railway infrastructure comes from cuttings, where the slope material directly correlates to the composition of surface sediments and geology. But also a large fraction of embankments in Great Britain are constructed from fill materials derived from plastic overconsolidated clays. This has led to a history of deformation problems requiring extensive maintenance and periodic slope failures (Perry *et al*., 1999), particularly noticeable in areas of high plasticity clays (Mott MacDonald, 2005).

The results of a detailed landslide and slope failure study by Pennington *et al*. (2014) indicate that there is a strong positive correlation between the probability of relatively small slope failures and statistically heavy rainfall events in the past. While such a correlation is less relevant for large landslides due to their dependence on a complex array of hydrogeological factors, these are minor landslides and slope failures that occur more often and cause the most rail infrastructure damage in the UK (Pennington *et al.*, 2014). Other climatic extremes such as high summer temperatures also have a detrimental impact on slope stability as they contribute to clay slope strength degradation caused by desiccation cracking (Loveridge *et al.*, 2010).

After having identified precipitation and soil clay content as important environmental variables behind slope failure probability, the availability and suitability of data was investigated and the following UK-wide datasets shortlisted that best capture environmental risk to railway lines:

- Met Office UKCP09 [http://ukclimateprojections.metoffice.gov.uk] climate dataset containing monthly long-term precipitation averages at 5×5km resolution for the 1961-1990 climate baseline period, used to describe 'normal' conditions in any given month.
- Met Office NIMROD [http://www.metoffice.gov.uk/industry/data/commercial/rainfall] live rainfall data at 5km spatial and 5-minute temporal resolution since 2004, obtained from satellite imagery, radar imagery and surface synoptic reports (Golding, 1998). This dataset is used to calculate deviations from 'normal' climate conditions and to identify extreme precipitation events.
- British Geological Survey [http://www.bgs.ac.uk/products/onshore/soilPMM.html] soil structure and parent material data at 1km resolution, used to identify areas with different levels of clay content in the soil. The percentage of clay content in soils was estimated using the BGS Parent Material database and soil texture triangle (Hodgson, 1997). As only clay was of interest, it was possible to split all soils into three groups using mean clay content: 77% (clay), 42% (clay to clayey loam, clay to sandy loam) and 24% (clayey loam to sandy loam, peaty clay, chalky clay to chalky loam).

Other data sources were found less suitable for automated risk prediction for a variety of reasons (Met Office station measurements – because of their point nature and potential interpolation problems, National Soil Map of England and Wales by Cranfield University – because of high licensing costs and difficulty of assessing soil clay content) and were not used in the implementation of the risk assessment tool described in this paper.

The complexity of soil saturation and slope failure processes cause certain levels of interaction inertia between the atmosphere and the pedosphere which is reasonably well captured by applying a 5-day sliding window over the environmental risk index and landslip observation events.

Figure 1 illustrates the spatial schema of environmental risk calculations from the three datasets described above. Risk assessed via this approach is not adjusted against infrastructure resilience or frequency of past events, but depends purely on the underlying geological conditions and precipitation history in any given location.

[FIGURE 1]

Precipitation time series, climate statistics, soil composition data and rail line geometries were all integrated into a geospatial decision support tool, presented in the following section.

3. Software architecture and implementation details

The interactive mapping tool described in this section (Figure 2) draws upon the cause-consequence relationships and data sources identified above and constitutes an online decision support system for visualisation of mathematical model results and ultimately the near real-time assessment of rail line risk across the UK.

[FIGURE 2]

Both the data sources and outputs are rasters, i.e. array of values arranged in a regular grid. The software is designed to meet two requirements:

1. Maintain a library of recent and historic rasters, to provide both near real-time presentations and to contextualize historic events in the datasets of land-slips and trees on the line.
2. Generate and present highly customised rasters on demand.

To accomplish these, a 24-hour cumulative rainfall raster was generated for each day from 2004 onwards, synchronised with the NIMROD database every 24 hours. These are generated by summing each day's rasters - one for each 5-minute interval (those that are not in NIMROD are interpolated). When added up, the original NIMROD units of $\frac{1}{32} \times \text{mm per hour}$ becomes:

$$\frac{1}{32} \times \text{mm per hour} \div 12 \text{ snapshots per hour} = \frac{1}{384} \times \text{mm}$$

In addition to the daily rainfall rasters, a default output raster is generated for each of those days. This is because calculating them on demand is relatively expensive computationally, each one typically taking a few seconds to generate.

Rasters are generated by a flexible script which accepts the following parameters:

1. Date - the final day rainfall to consider
2. List of coefficients to apply to the preceding days' rain (default: "1,1,1,1,1", i.e. 5 days, weighted equally)
3. Formula to apply (default "A/B*C*D"). Output rasters are integers, normalized to the range $[0,127]$

4. Threshold (default 150%). Regardless of the requested formula, an exceedance raster $\frac{A}{B}$ is always calculated, and 128 is added to all pixels where exceedance surpasses the threshold. In this way, a single raster can be used to present both exceedance and the formula outcome.

The rasters are presented as a new map layer, overlaid on top of the GeoSRM pilot map. In the map configuration, colours are assigned to each interval $I_n = [16n, 16n + 15]$. The colour assigned to I_{n+8} is a more intense or saturated version of the colour at I_n, since those two intervals represent the same range of formula values.

The actual values represented by each I_n are determined by the extreme values of each un-normalized raster, so these are retained, not only so that a key can be generated for each raster which includes the intervals' boundaries, but also so that it may be queried by users using the inspector tool (see Figure 2). For an explanation of other GeoSRM tools, see Sadler *et al.* (2016).

The variables within the formula are described in Table 1, below.

Table 1. Composition of the environmental risk formula

A	cumulative rainfall over the last n days, weighted according to coefficients $\mu_1, \mu_2, \dots, \mu_n$	$A = \left(\sum_{i \le n} \mu_i A_i\right) \div \left(384 \cdot \sum_{i \le n} \mu_i\right)$ mm per day in the unweighted case
B	monthly average rainfall	$B = \text{month's long-term average} \div \text{no. days in month}$ mm per day
C	soil type (proportion clay)	Percentage
D	track location (mask)	No units
E	GeoSRM derailment risk	FWI (Fatalities and Weighted Injuries per year)

It can be difficult to apply arbitrary formulae to rasters of differing resolutions in a manner which consistently generates meaningful results. The most reliable method identified was to use the GDAL (Geospatial Data Abstraction Library, a widely used open source tool for manipulating spatial data) raster calculation utility (gdal_calc.py), and manage source and intermediate rasters to meet gdal_calc's requirement that all input rasters be the same resolution and projection.

The NIMROD data and long-term average rainfall data are to 5km resolution; the soil type raster is to 1km; the rail network mask and GeoSRM derailment rasters need to be higher resolution in order to recognisably follow the track geometry, so are taken at 500m. To make calculations easier, the soil raster resolution was reduced to 5km, to match the rainfall rasters. The output of formulae that involve D or E make much more sense at a higher resolution, because it maintains the high level of spatial data granularity provided by GeoSRM derailment risk dataset. A constraint is therefore placed on the requested formula, that D or E, if included, should appear at the end, where it will be treated as a multiplicand to the entire previous calculation, e.g. a formula “A+B+CD” is interpreted as $(A + B + C) \cdot D$.

The computation is developed in stages (see Figure 3). First, A is calculated by adding the term from each A_i, one at a time. Then the requested formula (excluding any final D or E term) and exceedance rasters are computed. So far, all rasters have been at 5km resolution, but now resampled down to 500m, apply the D or E term if requested, and normalise the custom formula raster to integers in the range $[0,127]$. Finally, the normalized raster is combined with the exceedance raster (by gdal_calc.py with the formula “(E>threshold)*(N>0)*128+N” where N is the normalized raster, and E is exceedance).

[FIGURE 3]

This sequence is encoded within a single script (risk_calculator.sh), which also fetches and synchronises NIMROD data and calculates the daily cumulative rainfall rasters. It is executed to satisfy requests for new rasters, and also from a daily automatically scheduled task to keep the library of default rasters up to date.

Because each raster takes a few seconds to produce, the rasters are cached so that subsequent requests for recently requested rasters are timely - when the stored rasters start consuming too much filesystem space, the oldest (least recently accessed) are automatically removed.

Custom rasters are requested by web browser clients via an OGC compliant Web Mapping Service (WMS) (“GetMap”) request, passing date (mandatory) and formula, coefficients, threshold (optional) in the parameters (query string). Where a request for the raster overlay is detected, middleware code checks the cache and if necessary generates the raster, before appending its name to the request’s parameter list and passing it on to the mapping server.

While the backend supports generating rasters with all sorts of formulae, the present user interface only exposes the following 5, enabling users to work through the build-up from simple to more complex model calculations:

1. A (rainfall)
2. $A \div B$ (exceedance)
3. $A \div B \times C$ (exceedance on clay)
4. $A \div B \times C \times D$ (exceedance on clay, masked to the rail network)
5. $A \div B \times C \times E$ (GeoSRM derailment model results, modified by the environmental risk model)

For each of these formulae, controls are provided for manipulating the rainfall coefficients μ_i (which can be arbitrary or fitted to constant, linear or exponential curves), and for formulae 3, 4 and 5 a control is provided to adjust the relative impact (weighting) of exceedance vs. soil-type.

Mathematically, this could be done by applying an exponent to the sources, like $(A/B)^{\lambda_1} \times C^{\lambda_2}$, but this is discontinuous with respect to λ around zero (the soil raster in particular contains many zeros), so instead a balancing coefficient is used: $\left(1 + \lambda_1(A/B - 1)\right) \times \left(1 + \lambda_2(C - 1)\right)$, $\lambda_i \in [0,1]$.

Another interface control is provided to adjust the exceedance threshold, defaulting to 150% as recommended within other research (Pennington *et al.*, 2014) and a time control to select the period (any quarter from Q2 2004).

When a period is selected, any events that occurred during that time are displayed in two vector layers, one each for landslips and trees on the line, which can be individually toggled from the tool map layer control. Events which happen on or near that day are visually emphasised by size and with a highlighted outline.

The environmental risk tool is activated from within the GeoSRM pilot interface by clicking on the weather icon (in the main control cluster, Figure 2, top right). This causes:

- The most recent default raster to be overlaid on top of the GeoSRM
- The raster's key to appear (bottom right)
- Landslips and tree-fall layers to be added to the map "layers" control
- An inspector tool to be added to the control cluster
- The basic environmental controls to appear (bottom left)

The environmental controls consist of the drop-down list of periods, a slider control for days within that period, transport controls for nudging to the previous or next day, to the first or last day of a period, or for "animating" through each day, one after another. Buttons allow users to select each of the 5 basic formula types as listed above. When a raster is generated in response to an immediate request (as opposed to retrieval from a cache), the user is kept appraised of progress at the bottom of the control panel. Sequencing through custom rasters can be fairly slow (a few seconds each), but subsequent requests take advantage of the cache, so can play through much more smoothly.

As with other GeoSRM interface panels, the environmental control panel can be hidden to expose the map more fully. Also, they can be expanded to reveal more advanced controls, for:

- adjusting the number of days' rainfall and the coefficients to apply to them
- adjusting the balance between exceedance and soil-type
- modifying the exceedance threshold

When the ‘inspector’ mode is activated, clicking on a map pixel reveals a popup window explaining exactly how the value at that location was derived: the precise value from each source raster, and the formula that was applied.

4. Verification of results

To verify the performance of the suggested algorithm and the reliability of environmental risk prediction to GB rail infrastructure, the authors performed a series of statistical and GIS analysis tests on the resulting data. Due to the availability of data and involvement of the relevant train operators, the tests were run against a database of rail safety-relevant events across the country that occurred between January 2004 and February 2015. Within this period of time RSSB recorded a total of 680 landslips and 1214 tree-on-line events.

All events were georeferenced using track section geometries and event occurrence mileage, then compared to the results provided by the environmental risk prediction tool run using default parameters: 5-day cumulative precipitation floating window, 150% precipitation exceedance threshold and A/B*C*D formula as described in Section 3.

The fitness of this model was assessed by statistically comparing raster values near historic landslips with values from elsewhere along the rail network, as pixels that do not intersect with the rail network were not considered. Values that occurred at the location of a landslip event and up to five days before it, were considered “landslip” values. All other values are “no landslip” values. In the large dataset, which covers over a decade at 500m resolution, approximately 5000 values were found in the “landslip” set, and 150 million values in the “no landslip” set.

Table 2. Statistical evaluation of the results

	Landslip	No landslip
Count N	4,829	149,000,000
Mean μ	0.104	0.0590
Variance σ^2	0.0889	0.0582
Std. deviation σ	0.298	0.241

Both sets had distributions that were skewed heavily towards zero, with 87% of “no landslip” and well over 99% of “landslip” values falling in their respective bottom deciles.

Figure 4 shows the “landslip” set plotted against the overall average risk score of 0.028, and a corresponding map of locations with environmental risk score above the average.

[FIGURE 4]

This result is interpreted as follows: a majority of landslips occurred at times and locations that were not flagged by the environmental risk assessment algorithm and were most probably caused by reasons other than precipitation-induced clay and clayey soil saturation within the 5 day floating window. The results in Table 2 indicate that it is not yet possible to use the default raster to predict future landslip events with a high level of confidence. This might be because railway ground-works are engineered to be very resilient to environmental risk, so most unfavourable environmental episodes did not incur any damage. For an environmental risk model to score well, it would have to capture information that was unknown to the engineers who built and maintain the network.

5. Further work and expected benefits

A means of combining rasters is presented (in this case, precipitation and geology, but others could be introduced) according to arbitrary formulae. This method is also suitable for comparing competing environmental risk models, i.e., express each model as a formula, set of coefficients etc., use them to generate rasters, then compare them to the events dataset. Competing models can thereby be ranked, and their coefficients tuned.

So far, only the default raster settings have been analysed. In future work, selected models could be compared by randomly sampling a few days, saving time generating the rasters and gathering statistics from them. A simplistic score method would be to take a correlation coefficient ($r = 0.0011$ for the default raster set) and to search for model parameters with the highest correlation values

The methodology and architecture have been developed with potential further impacts and innovations in mind. There is great potential to incorporate 'live' incident data, more detailed asset data, a fuller range of real-time and historic environmental data, additional methodologies and formulae for calculating risk. Further development opportunities could include integration with sensors (e.g. Utterberry) to monitor land/vegetation stability, and integration of the tool and its output with other decision support systems.

Further improvement of this risk score can be achieved by incorporation of results of the relationship between the range of annual wetting/drying cycle and strain softening/slope failure in UK clays as investigated by Clarke & Smethurst (2014). To apply the results of this research, the authors suggest calculation of soil moisture deficit (SMD) values for every NIMROD-covered location by adopting the Gaussian-shaped potential evapotranspiration curve and then computing SMD with the real-time daily precipitation data. Computed for a year, this would yield the range of annual wetting/drying cycle over the last 365 days. The departure of this cycle from the location-specific climatically normal range calculated by analysis of the MORECS dataset such as appears in Clark (2002) would constitute an indication of an extreme wetting/drying cycle and hence increase the risk of slope failure, leading to derailments and damage to the rail infrastructure.

Another option to fine-tune the result would be to switch to shorter cumulative precipitation thresholds and focus on severe precipitation events only. A number of studies in the past (e.g. Stipanovic Oslakovic *et al.*, 2013) developed risk ranges and relationships between cumulative precipitation amounts and their impacts on the transport system, such as:

- Rainfall > 30 mm / 24 hours: Adverse impacts to the transport system may start to occur.
- 100 mm / 24 hours: Some adverse impacts are likely. Their severity depends on the resilience of the transport system.
- 150 mm / 24 hours: Weather phenomenon is so severe that it is virtually certain that some adverse impacts will occur. Coincidentally, the same precipitation amount of 150 mm has been reported in Freeborough *et al.* (2014) as the one correlating with the number and scale of landslides across the UK.

The already existing functionality to acquire and to analyse near-real-time precipitation data can potentially turn the tool into a useful instrument for operational monitoring and disaster preparedness purposes, flagging the areas at risk simultaneously with accumulation of rainfall at track sections.

Further development of the tool and successful validation of results will have measurable safety benefits. According to GB Rail Industry Safety Risk Model (v8.1), 0.39 fatalities and weighted injuries (FWI)/year result from landslips, while 0.08 FWI/year result from trees on the line events. By deploying such a system to prioritise preventative maintenance and to identify areas for improved monitoring, assuming ~10% effectiveness (not all landslips and trees on the line are due to rainfall), safety improvements could be achieved:

- 0.04 FWI/yr (saving the equivalent of a life every 25 years) for landslips
- 0.008 FWI/yr (saving the equivalent of a life every 125 years) for trees

Other than safety benefits, there are other potential positive impacts, e.g. performance (reduction in delay minutes, cancellations of services), reduction in knock-on damage to other infrastructure assets, e.g. track, track bed, signalling etc., greater efficiency in resource allocation/siting.

Benefits could be realised further by applying the same concept outside of the UK, depending on the availability of environmental data, or to the road network. Initial discussions with the Highways England have identified a similar base SRM model with potential to apply the same techniques.

The prototype was presented to a number of industry and academic stakeholders, identifying areas for potential development, possibly in partnership with industry. The research outcomes have identified clear potential benefits to Network Rail, RSSB and other industry stakeholders, all concerned with making the rail network resilient to environmental risk and targeting investment (and preventive measures) effectively for efficient management of infrastructure.

6. Conclusions

The authors have developed a scientifically robust methodology and an online interactive decision support system for identification of rail slope stability risk hotspots. This has established a way to flag certain locations for their near-real time vulnerability to environmental factors, with potential to contribute to further improvements in rail safety across Great Britain. This study demonstrates how historic and near real-time environmental raster data can be flexibly combined to enhance and extend the GeoSRM pilot, delivering spatially detailed and timely risk estimations while relying exclusively on freely available open data.

The methodology suggested was statistically evaluated using a nation-wide landslips event dataset. Further observation is still necessary to establish the best location-specific risk calculation formulae and thresholds. Changes to climate conditions and precipitation patterns that have been observed in the UK during the last decades may also require re-defining of a new climatic normality that might be different from the 1961-1990 long-term average.

The authors welcome comments and discussion regarding the choice of data and selection of the most appropriate risk assessment formulae, weightings and thresholds that would improve environmental risk prediction for GB rail infrastructure through the tools described in this paper.

Acknowledgements

The authors would like to acknowledge RSSB who commissioned and collaborated with the University of Southampton for the development of the original GeoSRM pilot, funded by the UK Department for Transport "T972 Piloting a geo-referenced safety risk model for the rail network in Great Britain".

Environmental extensions to GeoSRM, presented within this paper have been funded by NERC (Environmental Risks to Infrastructure Innovation Funding, grant NE/M008517/1) "A tool to improve prediction of real time environmental risk to UK rail infrastructure".

References

Chen P and Zhou J (2016) Effect of the built environment on automobile-involved pedestrian crash frequency and risk, Journal of Transport & Health 3 448-456.

Clark C (2002) Measured and estimated evaporation and soil moisture deficit in growers and the water industry. Meteorological Applications 9: 85–93.

Clarke D and Smethurst J (2014) Projections of climate change on the behaviour of clays in the UK. In, Unsaturated Soils: Research & Applications. UNSAT 2014: Unsaturated Soils: Research & Applications, CRC Press, Taylor and Francis: 191-196.

Dennis C and Somaiya K (2004) Development and Use of the UK Railway Network's Safety Risk Model, Probabilistic Safety Assessment and Management. PSAM 7 - ESREL '04, Berlin.

Dezman Z, de Andrade L et al. (2016) Hotspots and causes of motor vehicle crashes in Baltimore, Maryland: A geospatial analysis of five years of police crash and census data, Injury 47(11), 2450-2458.

Freeborough KA, Pennington CVL and Dashwood C (2014) *Landslides in the East Midlands*, Mercian Geologist 18(3): 185-187.

Gheorghe A, Birchmeier J, Vamanu D, Papazoglou I, Kröger W (2005) Comprehensive risk assessment for rail transportation of dangerous goods: a validated platform for decision support. Reliability Engineering & System Safety 88(3) 247-272.

Golding BW (1998) Nimrod: a system for generating automated very short range forecasts. Meteorological Applications 5(1): 1–16.

Hodgson, JM 1997 Soil Survey field handbook. Soil Survey Technical monograph No 5.Soil Survey and Land Research Centre, Silsoe, England.

Loveridge F, Spink T, Briggs K et al. (2010) The impact of climate and climate change on infrastructure slopes, with particular reference to southern England. Q J. Eng. Geol. Hydrogeol., 2010, 43 (4):461–472.

Mott MacDonald (2005) Seasonal preparedness, package 4: Statistics, Technical report, Mott MacDonald.

Pennington C, Dijkstra T, Lark M et al. (2014) Antecedent precipitation as a potential proxy for landslide incidence in South West United Kingdom. In: Sassa K, Canuti P, Yin Y, (Eds.) Landslide Science for a Safer Geoenvironment, Springer: 253–259.

Perry J, McGinnity B T and Russel D (1999) Railway earth structures: their condition, management and effect on track quality. Railway Engineering 1999, Second International Conference, London May 1999.

RSSB (Rail Safety and Standards Board) (2014) Safety risk model v 8.1. London, UK.

Sadler J, Griffin D, Gilchrist A et al. (2016) GeoSRM – Online geospatial safety risk model for the GB rail network. IET Intelligent Transport Systems 10 (1): 17–24.

Stipanovic Oslakovic I, Ter Maat H, Hartmann A and Dewulf G (2013) *Risk Assessment Of Climate Change Impacts On Railway Infrastructure,* Engineering Project Organization Conference, Working Paper Proceedings, http://www.epossociety.org/epoc2013/Papers/Stipanovic_terMaat_Hartmann_Dewulf.pdf

Wang X, Liu J, Khattak A, Clarke D (2016) Non-crossing rail-trespassing crashes in the past decade: A spatial approach to analysing injury severity, Safety Science 82, 44-55.

Figure and Table captions

Figure 1. Schema of the environmental risk calculation approach

Figure 2. Appearance and functionality of the GeoSRM environmental risk extension

Figure 3. Raster calculation flowchart

Figure 4. Spatial and numerical distribution of risk values at observed landslip events

Table 1. Composition of the environmental risk formula

Table 2. Statistical evaluation of the results