Enabling Resilient UK Energy Infrastructure: Natural Hazard Characterisation Technical Volumes and Case Studies

Volume 12:

Hazard Combinations

LC 0064_18V12

EDF ENERGY

energy technologies institute

IChemE ADVANCING CHEMICAL ENGINEERING WORLDWIDE

Institution of MECHANICAL ENGINEERS

Met Office

MOTT MACDONALD

Legal Statement

Author: Michael Sanderson (Met Office)
Chief Technical Officer: Hugo Winter (EDF Energy)

Version	Date	Details
0.1	02/03/18	Submitted for IPR
0.2	22/03/18	IPR comments addressed and submitted to CTO
1.0	10/05/18	CTO comments addressed and submitted to ETI
2.0	16/07/18	ETI and NHP3 Steering Committee comments addressed

Preface

This document forms part of the Energy Technologies Institute (ETI) project 'Low Carbon Electricity Generation Technologies: Review of Natural Hazards', funded by the ETI and led in delivery by the EDF Energy R&D UK Centre. The aim of the project has been to develop a consistent methodology for the characterisation of natural hazards, and to produce a high-quality peer-reviewed set of documents suitable for use across the energy industry to better understand the impact that natural hazards may have on new and existing infrastructure. This work is seen as vital given the drive to build new energy infrastructure and extend the life of current assets against the backdrop of increased exposure to a variety of natural hazards and the potential impact that climate change may have on the magnitude and frequency of these hazards.

The first edition of *Enabling Resilient UK Energy Infrastructure: Natural Hazard Characterisation Technical Volumes and Case Studies* has been funded by the ETI and authored by EDF Energy R&D UK Centre, with the Met Office and Mott MacDonald Limited. The ETI was active from 2007 to 2019, but to make the project outputs available to industry, organisations and individuals, the ETI has provided a licence to the Institution of Mechanical Engineers and Institution of Chemical Engineers to exploit the intellectual property. This enables these organisations to make these documents available and also update them as deemed appropriate.

The technical volumes outline the latest science in the field of natural hazard characterisation and are supported by case studies that illustrate how these approaches can be used to better understand the risks posed to UK infrastructure projects. The documents presented are split into a set of eleven technical volumes and five case studies.

Each technical volume aims to provide an overview of the latest science available to characterise the natural hazard under consideration within the specific volume. This includes a description of the phenomena related to a natural hazard, the data and methodologies that can be used to characterise the hazard, the regulatory context and emerging trends. These documents are aimed at the technical end-user with some prior knowledge of natural hazards and their potential impacts on infrastructure, who wishes to know more about the natural hazards and the methods that lie behind the values that are often quoted in guideline and standards documents. The volumes are not intended to be exhaustive and it is acknowledged that other approaches may be available to characterise a hazard. It has also not been the intention of the project to produce a set of standard engineering 'guidelines' (i.e. a step-by-step 'how to' guide for each hazard) since the specific hazards and levels of interest will vary widely depending on the infrastructure being built and where it is being built. For any energy-related projects affected by natural hazards, it is recommended that additional site- and infrastructure-specific analyses be undertaken by professionals. However, the approaches outlined

aim to provide a summary of methods available for each hazard across the energy industry. General advice on regulation and emerging trends are provided for each hazard as context, but again it is advised that end-users investigate in further detail for the latest developments relating to the hazard, technology, project and site of interest.

The case studies aim to illustrate how the approaches outlined in the technical volumes could be applied at a site to characterise a specific set of natural hazards. These documents are aimed at the less technical end-user who wants an illustration of the factors that need to be accounted for when characterising natural hazards at a site where there is new or existing infrastructure. The case studies have been chosen to illustrate several different locations around the UK with different types of site (e.g. offshore, onshore coastal site, onshore river site, etc.). Each of the natural hazards developed in the volumes has been illustrated for at least one of the case study locations. For the sake of expediency, only a small subset of all hazards has been illustrated at each site. However, it is noted that each case study site would require additional analysis for other natural hazards. Each case study should be seen as illustrative of the methods outlined in the technical volumes and the values derived at any site should not be directly used to provide site-specific values for any type of safety analysis. It is a project recommendation that detailed site-specific analysis should be undertaken by professionals when analysing the safety and operational performance of new or existing infrastructure. The case studies seek only to provide engineers and end-users with a better understanding of this type of analysis.

Whilst the requirements of specific legislation for a sub-sector of energy industry (e.g. nuclear, offshore) will take precedence, as outlined above, a more rounded understanding of hazard characterisation can be achieved by looking at the information provided in the technical volumes and case studies together. For the less technical end-user this may involve starting with a case study and then moving to the technical volume for additional detail, whereas the more technical end-user may jump straight to the volume and then cross-reference with the case study for an illustration of how to apply these methodologies at a specific site. The documents have been designed to fit together in either way and the choice is up to the end-user.

The documents should be referenced in the following way (examples given for a technical volume and case study):

ETI. 2018. *Enabling Resilient UK Energy Infrastructure: Natural Hazard Characterisation Technical Volumes and Case Studies,* Volume 1 — Introduction to the Technical Volumes and Case Studies. IMechE, IChemE.

ETI. 2018. *Enabling Resilient UK Energy Infrastructure: Natural Hazard Characterisation Technical Volumes and Case Studies,* Case Study 1 — Trawsfynydd. IMechE, IChemE.

Contents

Contents

Many natural hazards do not occur in isolation, but are often accompanied by other hazards. This situation is especially true for meteorological hazards. For example, low pressure weather systems bring winds and rain to the UK. Thunderstorms that produce damaging hailstones usually produce coincident heavy rain, and sometimes high winds. The impacts of high winds and heavy rain could be greater than the impacts of these hazards individually; for example, as well as flooding from the rainfall, high winds could force rain into assets containing electronics, causing damage to components, and short-circuits. Other hazards can be produced by separate phenomena but occur in rapid succession. If winter storms resulted in extensive flooding, and were followed by a prolonged period of very low temperatures, the flood waters could freeze over, exacerbating the impacts. Individual hazards are considered in detail in Volumes 2 to 11. This volume focuses on combinations of natural hazards, and specifically on ways to identify important and relevant combinations and methods to calculate joint probabilities of a given hazard combination occurring. The subject of hazard combinations is less advanced than methods for estimating probabilities of single hazards. At the time of writing, several methods for estimating joint probabilities are available, but no obvious overall method to tackle the problem has emerged.

Many different terms are used to describe combinations of hazards. Coincident hazards are those that occur at the same time, but need not be directly related. Dependent hazards are related, such as large hailstones accompanied by heavy rainfall, as both are produced by intense convective storms. A hazard that triggers another (e.g. persistent rainfall causing extensive flooding) is referred to as a 'cascade' or 'domino effect'. A hazard that increases the risk of another (e.g. a wildfire would remove vegetation and lead to an increased risk of a landslide during heavy rainfall) can be referred to as 'amplification'. The term 'multi-hazard' is also used when considering more than one hazard, but could refer to any of the terms discussed. In this technical volume, 'hazard combination' is used to refer to an event where two or more hazards have occurred at the same time or sufficiently close in time, resulting in an impact which could be greater than the sum of the component parts. Damage to an asset by one hazard could lead to an enhanced impact by a second. For example, if a roof was damaged by high winds, subsequent rainfall would be more hazardous if repairs to the roof were not completed.

For any assessment of combined hazards, due consideration should be given to the most relevant spatial scales for the asset being considered. For example, in Case Study 1 — Trawsfynydd, a point location is under study. Any hazards would need to be geographically coincident. Alternatively, if a large part of the whole transmission grid were considered, the

geographical definition would need to be much wider. The temporal constraints can be wider, over days or even weeks, depending on the types of hazard considered. For example, floods can occur over a wide range of timescales. Heavy rainfall can overwhelm drains resulting in temporary localised floods. More extensive and prolonged rain can produce riverine flooding as the floodwaters move downstream. These floods can occur hours or days after the rainfall.

Consecutive events of the same type are not considered here, but it is noted that they can cause significant disruption. Persistent rain from multiple storms was responsible for the 2013/2014 floods in the Somerset Levels. Some of the daily rainfall amounts were high, although none were exceptional.

The area of chalk geology that runs from the Chilterns to the Downs, north, west and south of London, gives rise to groundwater flooding when sufficient rainfall percolates into the chalk aquifer over time. This kind of flooding can occur weeks after the rain events, as in 2014 to the south and west of London. The flooding itself can last for days or weeks and in 2014 it affected road and rail networks, notably the Great Western Mainline near Maidenhead. Evolution of floods can therefore start with surface water flooding, followed by rivers bursting their banks as the surface water drains into the river system, with slow percolation into aquifers causing groundwater flooding if the geology and quantity/longevity conditions permit.

This project has identified 32 natural hazards of relevance to energy infrastructure. Not all combinations of these hazards are plausible. For physical reasons, some combinations are implausible. Other combinations may not be relevant for a location of interest. The hazards and combinations discussed here are meteorological in focus, but some of the methods described during the analyses could be used for other hazards.

Energy infrastructure, like many other sectors, is vulnerable to natural hazards including severe weather events. The combined impacts of certain combinations of hazards are not always straightforward to predict. For example, if strong winds were to accompany or follow heavy snowfall, the winds may enhance the impact (drifting snow blocking transport routes and building up against the sides of buildings), or reduce it (if the winds remove all of the snow except for the hardiest surface crust).

One example of combined hazards causing greater impacts than individual hazards is Storm Brian, which struck the UK on 21st October 2017. Very strong winds produced by this storm coincided with high tides, resulting in flooding of coastal areas in Wales and other parts of

the UK. Trains and ferries were cancelled and seafront roads closed. Further inland, fallen trees and other debris on railway tracks caused cancellations and disruption on some lines. This debris can also form dams on rivers, or block culverts and flow of water under bridges, causing flooding. Debris in rivers can also cause structural damage to infrastructure.

Hailstones, intense rainfall and lightning are produced by convective storms. Hail and rain often occur together or in quick succession; hail from a storm would either be followed by rain as the storm moves over a fixed point, or accompanied by rain. Flooding from heavy rainfall can be exacerbated by hail; the hail can block drains and channels, causing or enhancing flooding, as happened in Ludlow in June 1982 (*Meaden, 1982*) and Ottery St Mary in October 2008 (*Clark, 2011*). On 29th June 2012, a series of thunderstorms moved across the UK and produced heavy rain, large hailstones and high numbers of lightning strikes. Some contemporary reports showed images of roofs which were damaged by the hailstones, allowing rain to enter the building and cause further damage.

A succession of major storms during the winter of 2013/2014 produced extreme waves and swell which caused considerable erosion, damage and flooding in coastal areas of the UK (*Wadey et al., 2015*). These storms also caused power outages and disrupted travel owing to trees and other debris blocking rail lines and roads. September to December 2017 saw several deep low pressure systems which produced strong winds and, in some cases, heavy rain and flooding over the UK (Storm Aileen, 12th to 13th September; ex-hurricane Ophelia, 16th October; Storm Brian, 21st October; Storm Caroline, 7th to 8th December; Storm Dylan, 31st December). Flooding is considered in more detail in Volumes 5 — River Flooding, and 6 — Coastal Flooding.

Volcanic ash has a wide range of potential impacts on UK energy infrastructure, including flashover of high voltage power lines and other electrical equipment following deposition of ash, clogging of air inlets, and damage to computers. If large amounts of ash were washed out by rain, the wet ash would have the consistency of cement, and its weight could cause damage to infrastructure. Following the eruption of Mount Pinatubo in the Philippines in 1991, the ash was washed out by heavy rain, causing the roofs of many airport buildings in Manila to collapse (*Casadevall et al., 1996*). Wet ash could make road and other surfaces slippery. For example, an aircraft skidded off the runway at Manila airport because of reduced braking action on wet ash (*Casadevall et al., 1996*). Volcanic ash is considered in more detail in Volume 7 — Seismic, Volcanic and Geological Hazards.

Space weather refers to changes in the Sun's output or other phenomena, which can affect the Earth. Coronal mass ejections (CMEs) and solar flares can disrupt the Earth's magnetic field. Space weather events can also induce currents in electricity transmission grids, damage transformers and cause power blackouts. In March 1989, a solar storm caused a nine–hour outage of a power network in Quebec, when the induced currents in the transmission lines caused circuit breakers to trip. Changes in the solar cycle and the Sun's output have been shown to affect Europe's winter weather by modifying the circulation patterns (*Ineson et al., 2015*). Space weather is discussed in Volume 10.

Changes in the occurrence and magnitude of many individual hazards have been examined using observations and climate model simulations, but very few studies have considered changes in two or more hazards. For instance, there are studies of the impacts of climate change on heavy and extreme rainfall (*Kendon et al., 2014*) and hailstones (*Sanderson et al., 2015*) over the UK, but these studies did not assess concurrent changes in other hazards. One exception is the study by *Zappa et al. (2013)*, which examined projected changes in high winds and heavy rainfall for the UK and Europe from low pressure weather systems in a set of climate model simulations. Consequently, changes in many of the joint hazards under a warming climate have to be inferred from separate studies.

A large number of combinations of hazards are potentially relevant for the UK energy industry. Nine combinations considered the most important, based on industry needs and gaps in current scientific understanding, were identified at a workshop that included scientists and industry experts. These nine hazard combinations are listed in *Table 1*.

Table 1. List of nine hazard combinations identified in this project together with the most relevant meteorological phenomena.

Hazard combination	Convective storms (Summer)	Mid-latitude cyclones (all year)
Hail and wind	✓	
Hail and rain	✓	
Hail and warm temperatures	✓	
Hail and lightning	✓	✓
Wind and rain	✓	✓
Flooding followed by a large drop in temperature		✓
Storm surges combined with extreme waves and/or swell		✓
Volcanic ash dispersion and rain	✓	✓
Space weather and terrestrial weather	✓	✓

These combinations of hazards, and many others, are associated with one or both of low pressure systems and convective storms. Low pressure systems, also known as depressions or mid-latitude cyclones, are the most common type of weather experienced in the UK. They have spatial scales of the order of 1000 km, and are associated with clouds, rain and high winds. Low pressure systems generally form over the North Atlantic and travel eastwards, lasting for three to ten days. The exact paths taken by these systems are controlled by the jet stream. If the jet stream is positioned over the British Isles, the low pressure systems will bring rain, winds, possibly snow, and changeable weather. If the jet stream lies to the north of the UK, most of the country will experience warm and dry conditions, as the low pressure systems will travel to the north of Scotland. During winter, if the jet stream is positioned over the UK and then weakens and moves southwards, very cold easterly winds can blow over the UK, bringing very low temperatures, snow showers, blizzards, and possibly freezing rain.

Storm surges, swell and large waves are also associated with low pressure systems. A storm surge is an abnormal rise of water generated by the winds over and above the predicted astronomical tide. For the UK, swell is generated remotely by storms in the North Atlantic, and

propagates to the UK as low frequency waves. Swell carries a lot of energy and potential to cause damage when it reaches the shore. Large waves are produced by strong winds blowing over the surface of the sea; they have higher frequencies than swell. The size reached depends on the wind strength, the duration of the storm, and the fetch (the horizontal distance over which the wave-generating winds blow). The orientation of the coastline means the wind direction that produces high surges may not simultaneously produce large wave heights.

Convective storms, or thunderstorms, are very different to low pressure systems. They have much smaller spatial scales, of the order of a few kilometres, and tend to occur over localised areas. They are formed by rising air, which has been warmed at the surface. These types of storms can produce very heavy but localised rainfall, which in turn can cause flash-flooding and damage to property. They often produce lightning, and a small number of storms produce large hailstones. Some convective storms can produce short bursts of very high winds, and in extreme cases tornadoes.

The occurrence of ash from volcanic eruptions and space weather are fully independent of terrestrial conditions. These two hazards interact with convective storms and mid-latitude cyclones; ash is washed out by rain, and space weather can modulate lightning activity, which can be produced by both convective storms and mid-latitude cyclones. Ash is produced by volcanic eruptions, which are very difficult to predict. Any impact of the ash on the UK would depend on the strength of the eruption, the prevailing wind speeds and direction, and whether any low pressure systems were present at the time.

Space weather refers to three different phenomena: solar flares, solar energetic particles, and coronal mass ejections. If the particles have very high energies, they interact with the atmosphere, producing secondary showers of particles, which results in enhanced fluxes of neutrons at the surface. The state of microelectronic devices can be changed, thereby potentially affecting control and regulatory systems. The Earth's magnetic field can be modified by space weather, inducing currents in transmission lines, and potentially causing circuit breakers to trip, or damage to other equipment. Other studies have suggested that overall lightning activity is modulated by the polarity of the Sun's magnetic field and the solar wind (*Owens et al., 2014; Scott et al., 2014*).

Changes in the 11-year solar cycle have been shown to affect Europe's winter weather by modifying the circulation patterns (*Ineson et al., 2015*). Periods with low solar output tend to increase the chance of a 'blocking high pressure' in the Atlantic, resulting in colder, drier conditions across north-west Europe.

As the previous examples show, the relationship between the occurrence of a particular hazard combination and the occurrence of a particular impact (or impacts) is highly complex, and dependent on more than just the hazards in question. It is assumed that the reader will have some level of understanding about why a particular hazard combination is relevant for their purpose, which in turn is likely to stem from a recognition of the possible impact of that hazard combination on the energy infrastructure of interest.

Datasets describing many individual hazards are available. Some hazards (such as periods of high winds, heavy rain, or space weather events) would have to be defined and identified by users from individual records. The definition of 'high winds', for example, might vary depending on the type of asset under consideration. Information from different datasets would need to be combined to create a database of hazard combinations, which could then be used to assess the likelihood and potential severity of combinations of hazards. These issues are discussed further in this section.

3.1 Datasets of natural hazards for the UK

Datasets for the UK, which can be interrogated to investigate the hazards given in *Table 1*, are listed in *Table 2*. Many, but not all, of these datasets are freely available. Several historical flooding databases have been created from different sources, but it is unknown how consistent they are. No database of historical space weather events is known; they would have to be inferred by users from solar data. Some individual events have been described in other publications (for example, *Hapgood (2011)*).

Table 2. Summary of datasets for natural hazards in the UK. The second column 'Avail' indicates whether the data are freely available (Y = yes, N = no).

Type	Avail	Advantages	Disadvantages
UK weather station data (operated by the Met Office)[1]	N	Quality-controlled records of many weather variables (including wind, rain, air temperature).	Stations open and close; limited data periods for some stations; sparse coverage in some parts of the UK.
Hourly observations from weather stations: HadISD[2]	Y	Hourly values of several weather variables from thousands of stations worldwide; some records date back to 1935; constantly updated.	Data lengths and station opening/ closing times vary considerably; sparse coverage in some parts of the UK.
Daily observations from European weather stations: ECA&D[3]	Y	Quality controlled data from over 10,000 stations within Europe; updated to December 2017.	Coverage of UK sparse. Some data (e.g. wind speeds) not available over the UK.
The Tornado and Storm Research Organisation (also known as TORRO) hail database[4]	N	Comprehensive record of UK hailstorms, back to 1800 (and earlier).	Accuracy of some historical records unclear; only contains maximum hail-stone diameters; only records 'damaging' events.
Flooding maps[5]	Y	Maps of flooded areas in UK from 1946. Areas inundated can be calculated.	Map data may not be straightforward to analyse or interpret without specialist software.
Coastal flooding[6]	Y	Events from 1916 to present day.	Some events may be missing, especially before 1960. Actual areas inundated not available.
Inland and coastal flooding[7]	Y	Descriptions of historical flooding events in the UK, back as far as the 11th century.	Details of events vary considerably; actual areas inundated rarely mentioned.
Lightning (ATDnet operated by the Met Office)[8]	N	Accurate records of UK lightning locations and activity. Mostly detects cloud-to-ground strikes.	Lightning detection network stable from December 2007 only; other networks have longer records, and some have greater accuracy over the UK.
Ash deposition from volcanic eruptions[9]	Y	Historical events identified in peat and lake sediments, from 1000 A.D.	Identifying source of ash difficult. Amounts deposited not estimated.
Space weather events[10]	Y	Data series of many variables related to space weather, but does not list specific events.	Many different datasets from various satellites and other sources. Space weather events would have to be identified by a user.

1. https://www.metoffice.gov.uk/public/weather/climate-network/#?tab=climateNetwork.
2. https://www.metoffice.gov.uk/hadobs/hadisd/. See paper by Dunn et al. (2016) for a full description.
3. http://www.ecad.eu/
4. See paper by Webb et al. (2009) for a description of the hail database.
5. Maps available at: https://data.gov.uk/dataset/recorded-flood-outlines1
6. See paper by Haigh et al. (2017). Events can be visualised at: https://www.surgewatch.org/
7. See paper by Black and Law (2004). Searchable database at: http://www.cbhe.hydrology.org.uk/
8. https://www.metoffice.gov.uk/learning/making-a-forecast/first-steps/thunderstorms See Volume 9 — Lightning for a list of additional lightning datasets that include the UK.
9. See paper by Wastegård and Davies (2009) for more details.
10. Solar Data Analysis Center: https://umbra.nascom.nasa.gov/newsite/data.html

Several databases of historical flooding events have been created from a variety of data sources. A database of coastal flooding events was compiled from newspaper archives and tide gauge data by *Haigh et al. (2017)*. Monthly weather reports and UK climate summaries published by the Met Office, and monthly hydrological summaries published by the Centre for Ecology and Hydrology, were used to create a flooding database for the period 1884 to 2013 (*Stevens et al., 2016*). Another database of historical floods was based on reports from a wide variety of contemporary sources (*Black and Law, 2004*), and includes events dating back to the 11th century. It is not known how consistent these three databases are, or whether any attempt has been made to compare them.

3.2 Creating datasets of combinations of hazards

The datasets in *Table 2* describe single hazards, and would have to be combined to study combinations of hazards. For example, historical temperature records from weather stations could be analysed to see if any of the floods were accompanied by or followed closely by falls in temperature to below freezing. It would still be difficult to assess whether any of these floods subsequently froze over, or how thick the ice was if they did.

For other hazard combinations, creation of suitable datasets is more problematic. For example, observations of times and locations of hailstorms, together with measurements of the largest hailstone diameters, rely on human observers. The sizes of the hailstones may not always be recorded, and some hailstorms will have occurred far from any weather stations. Estimating concurrent wind speeds, temperatures and rainfall totals would therefore be difficult. The latter information could be inferred from archived weather forecast model data or rainfall radar data, if the time and location of the fall of hail were known accurately.

3.3 Dependence between hazards

The strength of the dependence between hazards may be influenced by the part of the distribution that is under consideration. For example, physical arguments suggest that rainfall and hailstones may be positively correlated — the more intense a convective storm, the higher the rainfall amounts and increased likelihood of large hailstones and/or large numbers of hailstones being produced. This was the case on 1st July 2015, when thunderstorms over parts of northern England and Scotland produced large hailstones and heavy rainfall concurrently. However, when dealing with extremes, the two hazards may be less well correlated or even independent. A storm in 1808 produced extremely large hailstones, up to 109 mm in diameter (*Clark, 2004*). The contemporary reports do not mention any rainfall (heavy or otherwise) accompanying this storm. Conversely, very intense pea-sized hail was produced by a storm which buried the town

of Ottery St Mary in up to 30 cm of hail in October 2008 (*Clark, 2011*). In this case, the hail was accompanied by intense rainfall that caused severe flooding of parts of the town. A similar event had occurred in Ludlow in 1982 (*Meaden, 1982*).

The association between some hazards may be dependent on the state of the atmosphere as quantified by large-scale indices (*Steptoe et al., 2018*). The pathways taken by low pressure systems over Europe are dependent on the sign and magnitude of the *North Atlantic Oscillation** (NAO). Another large-scale pressure pattern, the *Scandinavian Pattern* (SCP), modulates the activity of low pressure systems over northern Europe. A positive phase of the SCP corresponds to decreased occurrence of low pressure systems, causing them to be slowed down, weakened, and diverted away from the continent (*Steptoe et al., 2018*). Hence, any dependence between high winds and high rainfall may vary with the SCP.

*All technical terms marked in blue can be found in the Glossary section.

There are a wide range of potential hazards that could occur simultaneously or in rapid succession at any given site. When considering combinations of hazards, those that could not occur together should be screened out, together with those that are not relevant for a chosen site. In *Section 4.1*, examples of screening out of certain hazard combinations are described, together with suggestions for how certain combinations could be prioritised. Methods for the calculation of joint probabilities of two hazards (i.e. the probability of two hazards occurring simultaneously) are discussed briefly in *Section 4.2*. An example calculation of the joint probability of two meteorological hazards (high temperatures and low rainfall at Trawsfynydd) is given in *Section 4.3*.

4.1 Screening out and prioritisation of hazard combinations

Many hazard combinations are considered through design codes or standards (*ONR, 2017*). For example, high winds and heavy snowfall would increase the loading on a structure. The load from high winds could be combined with seismic stresses. Meteorological hazards considered within the *Nuclear Safety Technical Assessment Guide* are high winds; extreme drought; extreme air temperature; lightning; extreme hail, sleet or snow and icing; humidity; and climate change (which affects many of these hazards).

Many combinations of hazards can be screened out for a given site. Generally, this step involves eliminating those combinations that are unphysical or are incapable of posing a significant threat to the site, or are possible but highly unlikely (either at the site or elsewhere in the UK). Some hazard combinations are unphysical anywhere; meteorological observations indicate that many hazards could not occur together at the same location because they are either created by very different conditions or at very different times of year.

Some other hazard combinations can be neglected for certain sites. For example, hazards related to the marine environment, such as large swell, waves and very high sea levels are not relevant for inland sites. Other hazard combinations are possible for all sites, but either carry a low probability of occurrence, or at least one of the hazards is not predictable. In the latter case, any hazard combinations involving a meteorite impact or space weather event could occur, in principle, anywhere and at any time, making it difficult or impossible to conduct a quantitative assessment of such a combination.

For combinations involving any hazard which is possible but not predictable, one approach for the unpredictable hazard is to take a pre-defined scenario of occurrence for that hazard (where this exists), and consider the implications of that scenario for the energy infrastructure in question.

Depending on the nature of the other hazard in the combination, it may be possible to conduct a more quantitative analysis for the other hazard.

The characterisation of the remaining hazard combinations and their prioritisation is contingent on several considerations. A risk profile for each asset could be constructed, allowing different combinations of hazards to be prioritised. Questions to be asked include:

- What types of natural hazards are the site or infrastructure under study the most sensitive to?
- How critical are the possible impacts to the infrastructure, and are there data to characterise them?
- Are there maps of the frequency or severity of individual hazards that could be overlain, to help identify the most important combinations?
- Are certain combinations of hazards projected to increase in severity or frequency (or both) as the climate warms?

In a real assessment, there may be a need to characterise multiple combinations, even if they are considered to be low-probability, because low-probability hazards can often yield a major impact (hence their frequent appearance in some regulatory frameworks).

4.2 Methods for calculation of joint probabilities

A wide range of different methods are available for estimating the joint probability of two or more hazards. Here, the focus is on the joint probability of two hazards. Ideally, a large number of observed values would be used in order to provide a representative sample of the hazards and assess any dependence between them. A long series is particularly important if the focus is on extreme events which, by definition, are rare events. Use of a short data series means that the results could be strongly influenced by the chance occurrence of a small number of extreme events, or a lack of events. A data series could be extended with or replaced by modelled data, if the modelled data contain the same characteristics as the observations. Calculation of joint probabilities is a specialist area where expert guidance may be required. A brief overview of some methods is given here, and more details are given in *Appendix A*. An example calculation using *copulas* is given in *Section 4.3*.

- Independent — If two hazards are fully independent (e.g. a volcanic eruption on Iceland coinciding with north-westerly winds over the UK), their joint probability is simply the product of their individual probabilities.
- Empirically-based — If a very long series of observations of two hazards is available, their joint probability could be estimated by counting the number of days when the two

hazards occurred simultaneously, and dividing by the total number of days.

- Markov chains — If the probabilities of changing from one state to another are known (e.g. between days with and without hazards), the probability of a joint hazard occurring on the next day, or the third day ahead can be calculated.
- Statistical methods — Several different statistical methods have been used to estimate joint probabilities, including multivariate Bayesian frameworks, kernel density estimation, the conditional extremes model described by *Heffernan and Tawn (2004)* and the joint tails model of *Ledford and Tawn (1997)*. Copulas can also be used; these break down the joint probability calculation into two parts — construction of distribution functions for each marginal variable, and estimation of the dependence between them.

4.3 Example calculation of joint probabilities using copulas

4.3.1 Description of copulas

Copulas are statistical tools that describe the dependence between different variables (*Nelsen, 2010*). They allow a complex task, the specification of a joint *cumulative distribution function* (CDF), to be replaced by two simpler tasks: specification of a CDF for each variable, and specification of the dependence between the variables (see *Appendix A*). Copulas have been used in many areas, such as economics, finance and hydrology, to assess the joint probability of certain hazard combinations. In hydrology, copulas have been used to estimate *return periods* of droughts accompanied by high temperatures (*AghaKouchak et al., 2014*), and joint return periods of floods with large peak heights and volumes. Functions to fit copulas to data are readily available in the R programming language (*R Core Team, 2013*; *Yan, 2007*; see *Appendix B*).

There are many different types of copulas, which represent strong or weak dependence between variables. Some copulas are only valid for positively correlated variables (e.g. Clayton, Gumbel). Others may be used with moderately correlated variables but are not suitable for use with strongly correlated variables (e.g. the Farlie-Gumbel-Morgenstern copula). Selection of a copula is not necessarily straightforward. Some pragmatic choices can be made; for example, if the dependence between two variables were unknown, the selection of a copula that only permits positive correlations would not be appropriate. It is usually necessary to use other tools to estimate the strength and sign of dependence between two variables (e.g. the Kendall plot, described by *Genest and Boies (2003)*). Several different copulas would need to be tried and the best one(s) selected based on fits to the data by eye and goodness-of-fit statistics.

The chosen copula may not reproduce the correct dependence of the variables for very large extremes, where actual data are sparse or not available. Consequently, the joint probability of extreme values may be under - or over-estimated. Copulas have the tendency to scale quite poorly as more dimensions are introduced. If a full multivariate analysis of three or more variables was required, copulas can be very difficult to fit to the data. The user therefore must exercise some judgement and caution before deciding on which types of copula to use (*Dupuis, 2007*).

For the example demonstrated here, suitable weather observations are identified (*Section 4.3.2*). Next, the calculation of the joint probability with a copula is illustrated in five steps. These steps are summarised below and are described in more detail in *Sections 4.3.3* to *4.3.7*:

- assessment of possible dependencies;
- scaling the data so all values lie within the range 0 to 1;
- choice of and fitting of a copula;
- calculation of joint probabilities;
- estimation of uncertainty associated with the joint probabilities and return periods.

4.3.2 Weather observations

For any joint probability calculation, a long series of observations is required. A longer series of observations means more events would be available and confidence in the results would be higher. A weather station was operational at Trawsfynydd between 1961 and 1995, giving about 34 years of data. For this example, a longer series of data was obtained by selecting the nearest point of a 5 km daily gridded dataset created from UK weather stations (*Perry et al., 2009*). These gridded data were available for 1961 to 2016. They were used to calculate mean summer maximum temperatures and total summer rainfall (here, summer is June, July and August). The data series could be extended by pooling data from neighbouring grid points, assuming the climate at these grid points is similar to the site of interest (e.g. in altitude or exposure). Rainfall at surrounding points could, in theory, just as easily have fallen at the site, depending on the spatial scale considered.

4.3.3 Step 1: Assessing possible dependencies

The first step is to compare the two datasets and establish whether there is any evidence for dependence between them, and the nature of the dependence. The mean summer maximum temperatures and summer total rainfall amounts for each year are shown in *Figure 1(a)*, and the rainfall data are plotted against the temperature data in *Figure 1(b)*. There is little apparent dependence between the two datasets, except for high temperatures and low rainfall (<150 mm), where there is a suggestion of a negative correlation (high temperatures occur with low

rainfall). Some copulas are only suitable for positively correlated data. Hence, for this example, the temperature data are adjusted by subtracting them from 23 °C (all temperatures are less than 23 °C). By doing so, large temperatures are adjusted into small values and lower temperatures into large values. The summer rainfall totals are shown as a function of the adjusted temperatures in *Figure 1(c)*. Now, there is suggestion of a weak positive correlation when both variables have small values.

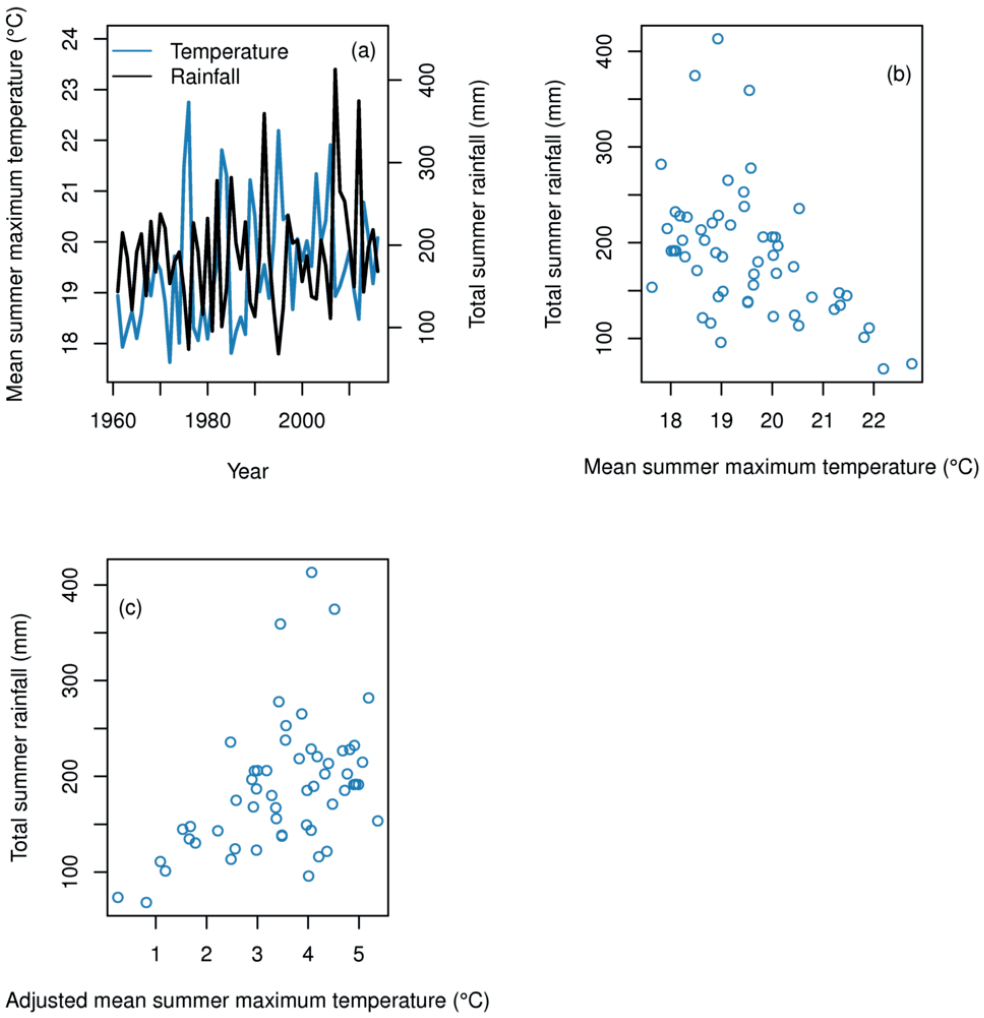

Figure 1. (a) Time series of mean summer maximum temperatures (blue) and summer total rainfall (black). (b) Summer rainfall as a function of summer temperature. (c) Summer rainfall as a function of adjusted temperatures.

Any dependence between two variables in the scatter plots in *Figure 1* can be hard to identify. Another useful tool for quickly assessing dependence (or independence) between two variables is the Kendall plot, or K-plot (*Genest and Boies, 2003*). Briefly, the amount of curvature in the graph is characteristic of the degree of dependence in the data. Different degrees of dependence between two variables are illustrated on a sample K-plot in *Figure 2(a)*. Variables that are completely independent lie along the leading diagonal (shown by the dashed line and black

crosses). Partial or perfect dependence are shown by the coloured lines marked a, b, c and d for variables that are positively and negatively correlated. Positively correlated variables lie above the diagonal, and negatively correlated variables lie below the diagonal.

A K-plot for the data shown in *Figure 1(c)* is illustrated in *Figure 2(b)*. The data points lie between the leading diagonal and the curved dotted line, indicating a partial positive dependence between the temperatures and rainfall, as was inferred from *Figure 1(c)*.

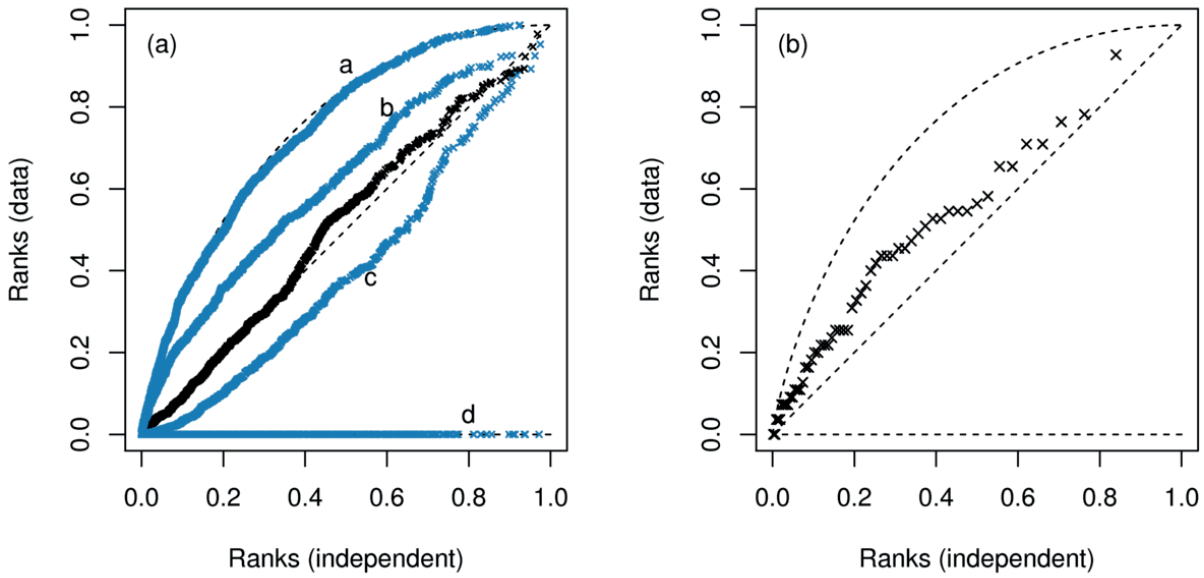

Figure 2. (a) Example of K-plot, showing the appearance of two datasets with no dependence (black line on leading diagonal), perfect positive dependence (curve a), partial positive dependence (curve b), partial negative dependence (curve c), perfect negative dependence (curve d). (b) K-plot for summer rainfall and adjusted temperatures at Trawsfynydd. Note that the ranks have been normalised to the range [0, 1] by dividing by the number of data points.

4.3.4 Step 2: Transformation of the data to the range 0 to 1

Before fitting a copula, the data have to be transformed into the range 0 to 1. One method for transforming the data into the range 0 to 1 is to fit a function to each of the two datasets, which describes their distributions. In the context of copulas, this function is called the *marginal distribution function*. A marginal distribution function describes the probability distribution of a single variable (Y). Integration of this function between minus infinity and a value y produces the CDF. The CDF evaluated at a value y is the probability that the variable Y will have a value less than or equal to y. All probabilities have values between 0 and 1.

Several mathematical functions were fitted to the temperature and rainfall data and the function with the best fit was selected based on *Akaike information criterion* (AIC) and other goodness-of-fit statistics. For the data used here, a Weibull function was selected for the

temperature data (*Figure 3*) and a gamma function for the rainfall data (*Figure 4*). Gamma functions are commonly used to describe rainfall distributions (*Hanson and Vogel, 2008*).

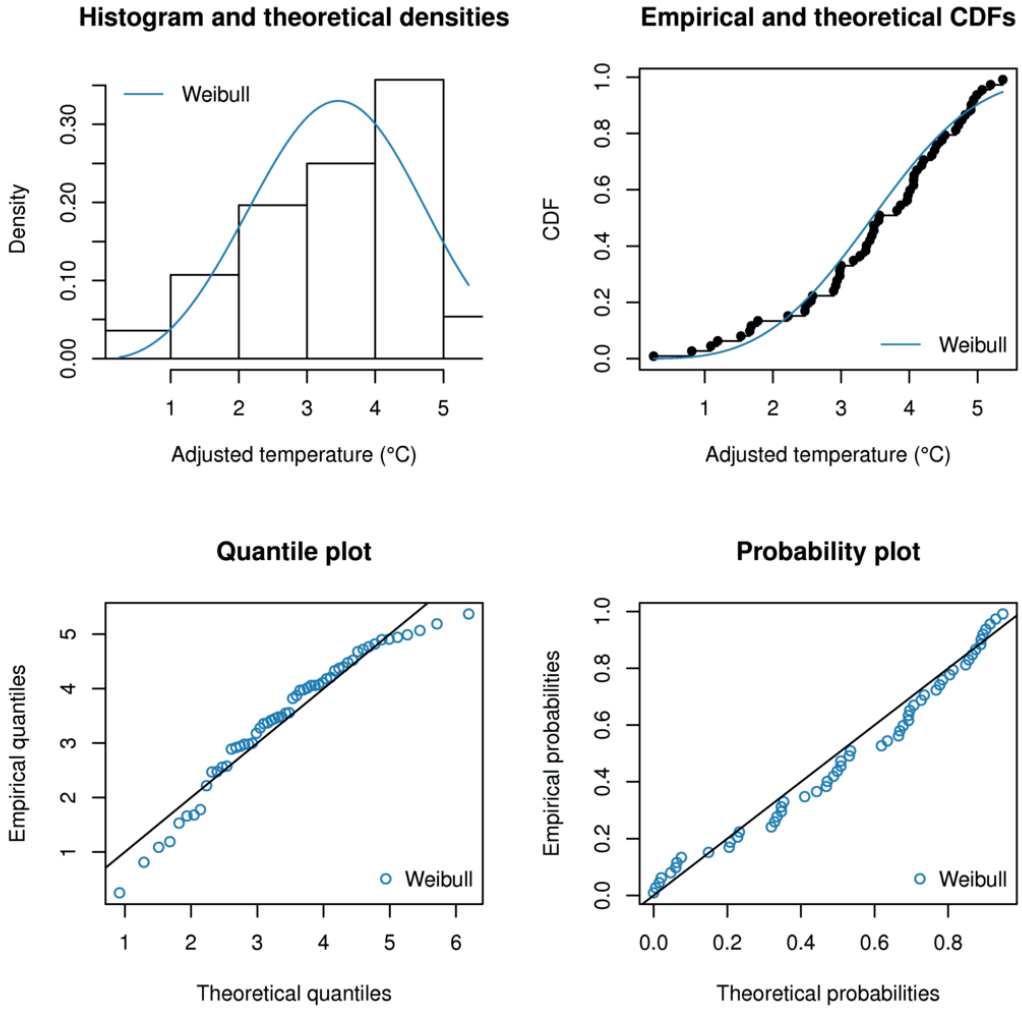

Figure 3. Fitting of a Weibull distribution to the transformed temperature data. The distributions of the data and fitted function (densities) are shown in the top left panel. The empirical (derived from the data) and theoretical (from the fitted Weibull function) cumulative distribution functions are shown in the upper right panel. The quantile plot (lower left panel) compares the quantiles of the data distribution with those of the Weibull function. The probability plot (lower right panel) compares the empirical cumulative distribution function of the data with the cumulative distribution function from the Weibull distribution. The quantile plot highlights differences in the tails of the distributions, whereas the probability plot highlights differences in the centres of the distributions.

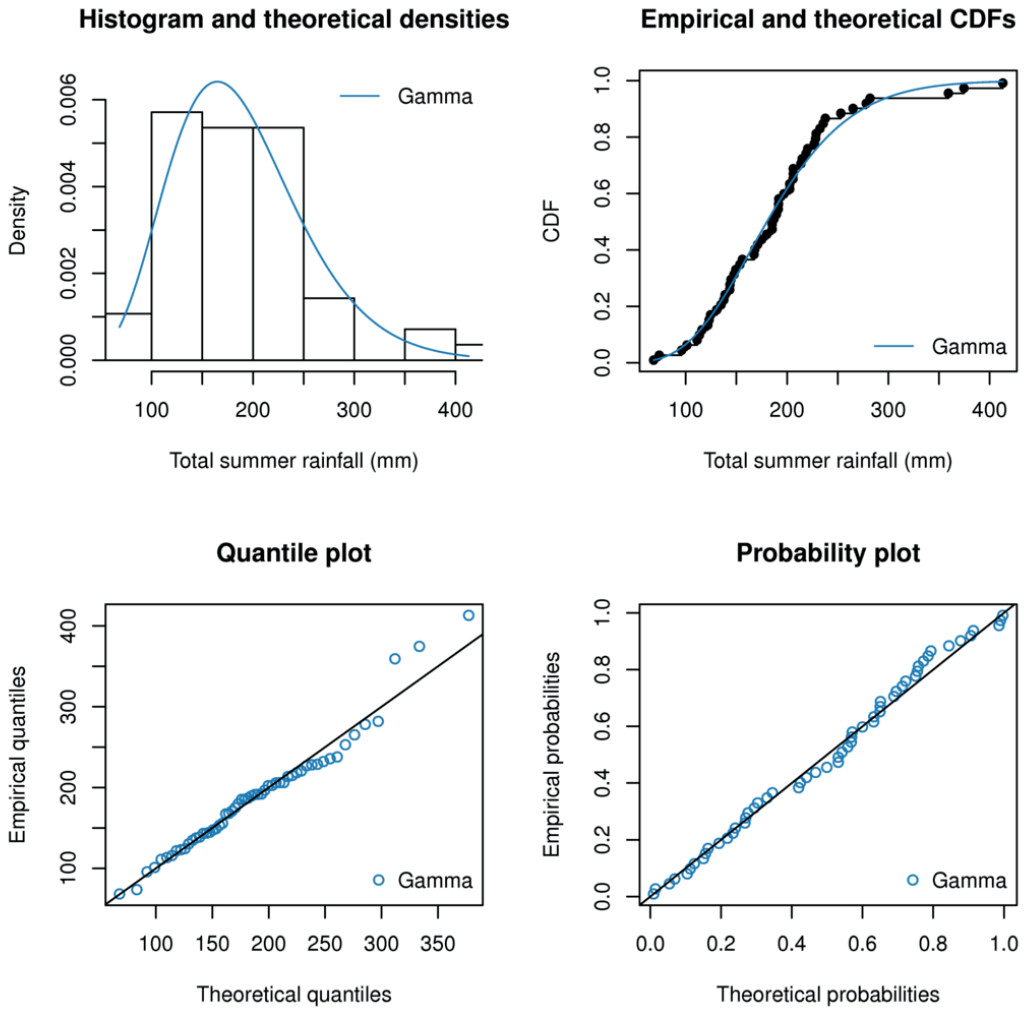

Figure 4. Fitting of a gamma distribution to the rainfall data. The distributions of the data and fitted gamma function (densities) are shown in the top left panel. The empirical (derived from the data) and theoretical (from the fitted gamma function) cumulative distribution functions are shown in the upper right panel. The quantile plot (lower left panel) compares the quantiles of the data distribution with those of the gamma function. The probability plot (lower right panel) compares the empirical cumulative distribution function of the data with the cumulative distribution function from the gamma distribution. The quantile plot highlights differences in the tails of the distributions, whereas the probability plot highlights differences in the centres of the distributions.

In *Figure 3*, the probability plot shows that the Weibull function reproduces the middle of the distribution reasonably well, but the quantile plot indicates that the tails of the distribution are not captured very well. In *Figure 4*, all four panels indicate that the gamma function fits the data reasonably well, except for the very largest rainfall values, which are underestimated. Better fits to the distributions could possibly be found if other functions were tested. Alternatively, a better fit might be obtained if a Generalised Pareto Distribution (GPD) was used for data above some threshold and an empirical distribution function below the threshold (*Heffernan and Tawn, 2004*).

The marginal distribution functions (top left panels of *Figure 3* and *Figure 4*) were used to calculate the corresponding CDFs (top right panels of *Figure 3* and *Figure 4*). The transformation of the data to the range 0 to 1 is achieved by first locating each the data value on the x-axis of the CDF figure, and then reading off the corresponding probability on the y-axis.

4.3.5 Step 3: Choosing and fitting a copula

Now that the marginal distributions have been estimated, a copula needs to be selected, which represents the dependence between the two variables. There are many different copula functions that represent various dependencies of two datasets — weak or strong dependence, dependence only for very large or very small values, or dependencies for extreme high and extreme low values of both variables (*Dupuis, 2007*; *Yan, 2007*). The data shown in *Figure 1(c)* indicates dependence when both variables have low values, but no dependence at higher values, and that the dependence is positive. The Clayton copula is suitable for this type of dependence; other copulas were tried but gave poorer fits to the data, based on AIC and visual comparisons of scatter plots of the simulated and observed data. *Genest et al. (2009)* reviewed methods for assessing fits by copulas, and a practical example of choosing a copula is given by *Requena et al. (2013)*.

The Clayton copula was fitted to the transformed data (i.e. the temperature and rainfall data transformed to the range 0 to 1 using the marginal distributions identified above). Next, 1000 values were randomly sampled from the fitted Clayton copula, which are also in the range 0 to 1. These random values are transformed back to the original data (i.e. actual values of temperature and rainfall) using the inverses of the CDFs shown in *Figure 3* and *Figure 4*. A scatter plot of the 1000 values generated from the Clayton copula and the observed data are shown in *Figure 5*. The simulated values have a similar pattern to the observed data, suggesting that the Clayton copula was a reasonable choice. Some of the simulated values lie outside of the range of the observed data. Correlations (based on *Kendall's tau*) in the original and simulated data are 0.34 and 0.27 respectively. These values indicate weak correlation, as before (a value of one for Kendall's tau indicates a perfect correlation and a value of zero indicates no correlation); the two correlations are in reasonable agreement, but ideally would be closer in value.

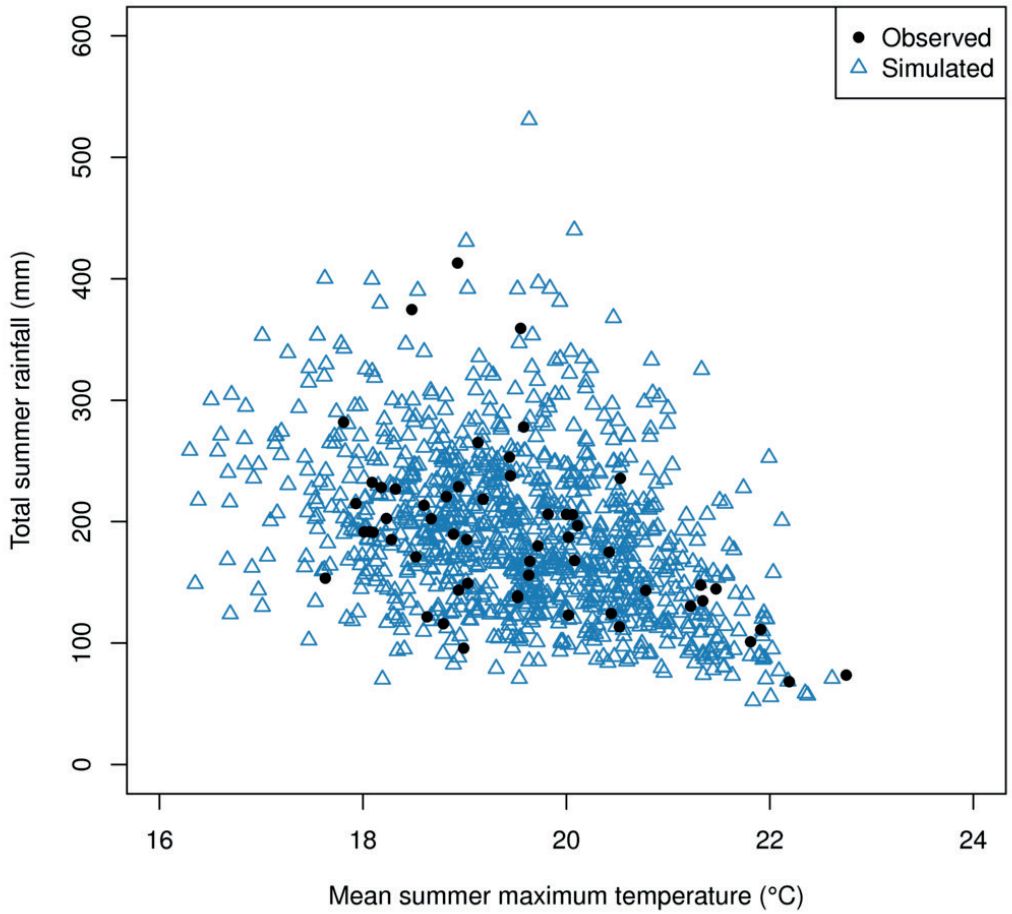

Figure 5. Scatter plot of observed data (solid circles) and simulated data (open triangles) using the Clayton copula.

4.3.6 Step 4: Calculation of joint probabilities

The joint probability of an event can now be calculated using the Clayton copula. For this example, the probability of summer mean temperatures exceeding 22 °C and summer rainfall being below 100 mm at Trawsfynydd will be calculated, corresponding to very warm and dry conditions. Recall that the temperature data were adjusted by subtracting them from 23 °C, so the calculation uses the condition that the adjusted temperature is less than or equal to 1.0 °C. First, a large number of random values (100,000) are generated using the fitted copula, which are then transformed back to actual data values using the inverses of the CDFs. Next, the joint probability is calculated as the number of randomly generated points for which the temperature is less than 1 °C and the rainfall is less than 100 mm, divided by the total number of points generated. The return period is the inverse of the probability multiplied by the time period that the observations correspond to. For the condition used here, 869 points met the two conditions above, so that the annual probability of the event is 0.00869, and the return period is 115 years.

4.3.7 Step 5: Estimation of uncertainty

The uncertainty in the joint probability and return periods calculated with the copula can be estimated in two ways. First, given that the interest is in rare events, the probability of such an event occurring is low. If it is also assumed that the number of randomly-generated samples which meet the two extreme criteria follows a Poisson distribution (which is a reasonable assumption), the uncertainty in the number of samples, *n*, is the square root of *n*. Hence, for this example, with a single randomly generated sample of 100,000 points:

- Number of samples meeting both thresholds = 869. Uncertainty = $\sqrt{869} = 30$.
- Probability = 869/100,000 = 0.00869.
- Uncertainty in probability = $(1/100{,}000) \times \sqrt{869} = 0.00030$.
- Return period = reciprocal of probability = 115 ± 4 years.

An alternative method for estimating the uncertainty in the probability is to generate random samples multiple times, and produce large numbers of estimates of counts above the thresholds. The mean number of counts and the standard deviation can then be calculated. As an example, 20,000 sets of 100,000 samples were generated, and the number of samples in each set for which both thresholds were exceeded was noted. This process produced 20,000 estimates of the number of times both thresholds were exceeded. The mean and standard deviation of these estimates were 853 and 29 respectively, which are very similar to the values of 869 and 30 calculated above. The corresponding return period is 117 ± 4 years.

Looking at the observed data (*Figure 1* and *Figure 5*), there are in fact two data points that fulfil the joint criteria above. In a data series spanning approximately 50 years, an approximate return period could be inferred of 25 years. The copula method described above effectively 'fills in' the dataset with many more points (as represented in *Figure 5*), and allows a return period to be estimated with smaller uncertainty — based on this larger pseudo-dataset — that is longer than that based on the observed data alone. This result suggests that the two observed points are quite unusual despite their having occurred within the same ~50-year period. The difference between the return periods estimated in these different ways indicates one of the potential pitfalls of basing such a calculation solely on observed data. The differences could also be due to the copula not fitting the distributions very well.

If it is assumed that the two events (low rainfall and high temperatures) are independent, the individual probabilities of low rainfall and high temperature can be estimated from the CDFs in *Figure 3* and *Figure 4* and then multiplied to estimate the joint probability (*Section 4.2*).

Under these assumptions, the joint probability is very small, 0.00076, corresponding to a return period of over 1300 years.

Commonly used approaches for univariate risk estimation significantly underestimate or overestimate the return period. Based on Weibull's method, the univariate return periods of the summer rainfall amounts at Trawsfynydd can be estimated based on their ranks (*AghaKouchak et al., 2014*). A rainfall amount of 100 mm has a return period of about 14 years, which is much smaller (i.e. more frequent) than the return period for concurrent low rainfall and high temperatures (115 years). Overall, these different return period estimates illustrate the importance of accounting for dependence when estimating joint probabilities of concurrent extremes.

4.4 Conclusion

Natural hazards are often considered in isolation. However, many natural hazards (particularly meteorological hazards) occur simultaneously or in rapid succession, and so enhance the impacts of an existing situation. Some combinations may be screened out because they are unphysical due to their nature, or would not pose a significant threat. For some sites, certain combinations will not be relevant — maritime hazards can be ignored for inland sites, for example. Some other hazard combinations are possible but are very difficult to predict — but this does not mean they can be ignored.

This assessment of hazard combinations has looked at pairs of natural hazards and a range of methods for calculating their joint probabilities, each of which has advantages and disadvantages. One realistic hazard combination — low summer rainfall coinciding with high temperatures — was analysed in detail to provide an example of how to calculate the joint probability of a hazard combination. This illustrative calculation relied on the use of a copula to represent the dependence between low summer rainfall and high temperatures. The joint return period of mean summer maximum temperatures exceeding 22 °C and summer rainfall being below 100 mm at Trawsfynydd was calculated to be 115 years. Methods for estimating probabilities and return periods of combinations of hazards are less well developed than those for single hazards. Several different methods for estimating joint probabilities are available, but no particular method has emerged as a definitive solution.

5. Related phenomena

There are no particular minor phenomena associated with hazard combinations as outlined in this technical volume.

Combinations of natural hazards have received very little attention in regulations until recently. As a result of the Fukushima event, the Office of Nuclear Regulation is in the process of updating its Technical Assessment Guides (TAGs). The following expectations with regards to consideration of combined hazards are anticipated from the TAG 13 update:

- Licensees should take into account combinations of External Hazards that could reasonably be expected to occur at a given site. Combinations of hazards should be identified and considered as part of Design Basis Analysis, Probabilistic Safety Analysis and Severe Accident Analysis.
- Licensees should follow a systematic process to identify and categorise hazard combinations and should then screen those hazards on the basis of plant effects and occurrence frequency.

7.1 High winds

Strong winds and heavy precipitation over the UK are produced by low pressure weather systems (also known as depressions or extratropical cyclones). *Zappa et al. (2013)* analysed changes in numbers of low pressure systems and associated wind speeds and rainfall between the end of the 20th century and the 2080s, as simulated by a large set of global climate models. One clear response from the models was an increase in rainfall intensity consistent with the warming of surface air temperature. In the multi-model mean, a 26% (winter) and 33% (summer) increase was projected in the number of cyclones with precipitation intensity within the top 10% of the historical simulations.

The mean wind intensity related to low pressure systems is projected to decrease, while the average wind intensity related to the deepest systems remained mostly unchanged in the North Atlantic (*Zappa et al., 2013*). A small increase in the number of cyclones (3%) and associated wind speeds (3%) was projected over the UK during winter by the 2080s. In the summer months, little change in numbers and wind speeds was projected. This analysis suggests that the probability of high wind speeds accompanied by heavy rainfall could increase slightly under a warming climate during winter.

7.2 Large waves

Aarnes et al. (2017) investigated annual changes in significant wave heights in the North Atlantic by the end of the 21st century using simulated winds from six global climate models to drive a wave model. They found that annual mean wave heights were expected to decrease in the future, and the largest decreases were simulated under a high emissions scenario. This decrease results from the projected reduction in mean wind speeds associated with low pressure systems (*Zappa et al., 2013*). However, increases in the heights of the most extreme waves were projected along the western coast of the UK.

7.3 Heavy rainfall

Kendon et al. (2014) studied heavy rainfall from convective storms over the UK using a very high resolution (1.5 km) climate model. They showed that the numbers of storms and the rainfall from these storms were projected to increase during the summer months under a warming climate. However, any changes in wind gusts or other hazards such as hail were not considered. The single study of the effects of climate change on hailstones over the UK by *Sanderson et al. (2015)* suggested that large hail would decrease in frequency. However, these authors used a lower resolution climate model (25 km) which could not resolve convective storms directly. They did not examine simultaneous changes in other hazards.

7.4 Volcanic ash

Any effect of climate change on volcanic ash deposition over the UK is very uncertain. The simulated increases in rainfall from low pressure systems noted by *Zappa et al. (2013)* could have two different and opposite effects on the ash. If an eruption coincided with the presence of a low pressure system over the volcano, more of the ash could be washed out before it reached the UK. However, if the ash was transported to the UK under drier conditions, and was then washed out, more of the ash could be washed out, potentially increasing the risk of this hazard.

7.5 Space weather

The occurrence of a space weather event is independent of the climatic conditions on Earth. Some studies have suggested that lightning activity is modulated by the solar wind and polarity of the Sun's magnetic field (*Scott et al., 2014*; *Owens et al., 2014*). If lightning flash rates do increase as the climate warms, as suggested by many studies, the flash rates could be enhanced further by space weather events.

7.6 Forthcoming climate projections

A new set of climate projections for the UK, called UKCP18, is currently under development (*UKCP Project, 2018*). All project outputs will be released during November 2018, including very high resolution (2.2 km) climate projections, which are intended to provide information on high impact events such as localised heavy rainfall during summer. These projections will use one of the newer emissions scenarios called Representative Concentration Pathways (RCPs).

References

Aarnes OJ, Reistad M, Breivik Ø, Bitner-Gregersen E, Ingolf Eide L, Gramstad O, Magnusson AK, Natvig B, Vanem E. 2017. Projected changes in significant wave height toward the end of the 21st century: Northeast Atlantic. *Journal of Geophysical Research: Oceans,* 122, 3394–3403. *doi: 10.1002/2016JC012521*

AghaKouchak A, Cheng L, Mazdiyasni O, Farahmand A. 2014. Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. *Geophysical Research Letters,* 41, 8847–8852. *doi: 10.1002/2014GL062308*

Black AR, Law FM. 2004. Development and utilization of a national web-based chronology of hydrological events. *Hydrological Sciences Journal,* 49, 237–246. *doi: 10.1623/hysj.49.2.237.34835*

Casadevall TJ, Delos Reyes PJ, Schneider DJ. 1996. The 1991 Pinatubo eruptions and their effects on aircraft operations. In *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines,* 625–636.

Clark C. 2004. The heatwave over England and the great hailstorm in Somerset, July 1808. *Weather,* 59, 172–176. *doi: 10.1256/wea.04.04*

Clark MR. 2011. An observational study of the exceptional 'Ottery St Mary' thunderstorm of 30 October 2008. *Meteorological Applications,* 18, 137–154. *doi: 10.1002/met.187*

Dunn RJH, Willett KM, Parker DE, Mitchell L. 2016. Expanding HadISD: quality-controlled, sub-daily station data from 1931. *Geoscientific Instrumentation, Methods and Data Systems,* 5, 473–491. *doi: 10.5194/gi-5-473-2016*

Dupuis DJ. 2007. Using copulas in hydrology: Benefits, cautions, and issues. *Journal of Hydrologic Engineering,* 12, 381–393. *doi: 10.1061/(ASCE)1084-0699(2007)12:4(381)*

Eastoe E, Koukoulas S, Jonathan P. 2013. Statistical measures of extremal dependence illustrated using measured sea surface elevations from a neighbourhood of coastal locations. *Ocean Engineering,* 62, 68–77. *doi: 10.1016/j.oceaneng.2013.01.002*

Genest C, Boies J-C. 2003. Detecting dependence with Kendall plots. *The American Statistician,* 57, 275–284. doi: 10.1198/0003130032431

Genest C, Rémillard B, Beaudoin D. 2009. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics,* 44, 199–213. doi: 10.1016/j.insmatheco.2007.10.005

Haigh ID, Ozsoy O, Wadey MP, Nicholls RJ, Gallop SL, Wahl T, Brown JM. 2017. An improved database of coastal flooding in the United Kingdom from 1915 to 2016. *Scientific Data,* 4, 170100. doi:10.1038/sdata.2017.100

Hanson LS, Vogel, R. 2008. The probability distribution of daily rainfall in the United States. In *World Environmental and Water Resources Congress 2008: Ahupua'A,* 1–10. doi:10.1061/40976(316)585

Hapgood M. 2011. Towards a scientific understanding of the risk from extreme space weather. A*dvances in Space Weather,* 47, 2059–2072. doi: 10.1016/j.asr.2010.02.007

Heffernan JE, Tawn JA. 2004. A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 66, 497–546. doi: 10.1111/j.1467-9868.2004.02050.x

Ineson S, Maycock AC, Gray LJ, Scaife AA, Dunstone NJ, Harder JW, Knight JR, Lockwood M, Manners JC, Wood RA. 2015. Regional climate impacts of a possible future grand solar minimum. *Nature Communications,* 6, 7535, doi: 10.1038/ncomms8535

Keef C, Papastathopoulos I, Tawn, JA. 2013. Estimation of the conditional distribution of a vector variable given that one of its components is large: additional constraints for the Heffernan and Tawn model. *Journal of Multivariate Analysis,* 115, 396–404. doi: 10.1016/j.jmva.2012.10.012

Kendon EJ, Roberts NM, Fowler HJ, Roberts MJ, Chan SC. 2014. Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change,* 4, 570–576. doi: 10.1038/nclimate2258

Ledford AW, Tawn JA. 1997. Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 59, 475–499. *doi: 10.1111/1467-9868.00080*

Meaden GT. 1982. Hailstorms in north-east Bristol, 18 June 1982, and Ludlow, 26 June 1982. *Journal of Meteorology,* 7, 224–225.

Murphy JM, Sexton DMH, Jenkins GJ, Boorman PM, Booth BBB, Brown CC, Clark RT, Collins M, Harris GR, Kendon EJ, Betts RA, Brown SJ, Howard TP, Humphrey KA, McCarthy MP, McDonald RE, Stephens A, Wallace C, Warren R, Wilby R, Wood RA. 2009. *UK Climate Projections Science Report: Climate change projections.* Met Office Hadley Centre, Exeter, UK. Available at: *http://ukclimateprojections.metoffice.gov.uk/media.jsp?mediaid=87893* (accessed on 10th May 2018).

Nelsen RB. 2010. *An Introduction to Copulas.* Springer Series in Statistics, Second Edition, Springer-Verlag, New York, US.

ONR. 2017. *Nuclear Safety Technical Assessment Guide*, NS-TAST-GD-013 (Revision 6), *http://www.onr.org.uk/operational/tech_asst_guides/* (accessed on 10th May 2018).

Owens MJ, Scott CJ, Lockwood M, Barnard L, Harrison RG, Nicoll K, Watt C, Bennett AJ. 2014. Modulation of UK lightning by heliospheric magnetic field polarity. *Environmental Research Letters,* 9, 115009. *doi: 10.1088/1748-9326/9/11/115009*

Perry M, Hollis D, Elms M. 2009. *The Generation of Daily Gridded Datasets of Temperature and Rainfall for the UK.* National Climate Information Centre, Climate Memorandum No. 24, Met Office, UK.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. Available at: *http://www.R-project.org/* (accessed on 10th May 2018).

Requena AI, Mediero L, Garrote L. 2013. A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation. *Hydrology and Earth System Science,* 17, 3023–3038. *doi: 10.5194/hess-17-3023-2013*

Sammonds P, McGuire B, Edwards S. (Eds). 2010. *Volcanic Hazard from Iceland: Analysis and Implications of the Eyjafjallajökull Eruption.* UCL Institute for Risk and Disaster Reduction, London, UK.

Sanderson MG, Hand WH, Groenemeijer P, Boorman PM, Webb JDC, McColl LJ. 2015. Projected changes in hailstorms during the 21st century over the United Kingdom. *International Journal of Climatology,* 35, 15–24. *doi: 10.1002/joc.3958*

Santhrosh D, Srinivas W. 2013. Bivariate frequency analysis of floods using a diffusion based kernel density estimator. *Water Resources Research,* 49, 8328–8343. *doi: 10.1002/2011WR010777*

Scott CJ, Harrison RG, Owens MJ, Lockwood M, Barnard L. 2014. Evidence for solar wind modulation of lightning. *Environmental Research Letters,* 9, 055004. *doi: 10.1088/1748-9326/9/5/055004*

Sexton DMH, Murphy JM, Collins M, Webb MJ. 2011. Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. *Climate Dynamics,* 38, 2513–2542. *doi: 10.1007/s00382-011-1208-9*

Steptoe H, Jones SEO, Fox H. 2018. Correlations between extreme atmospheric hazards and global teleconnections: Implications for multihazard resilience. *Reviews of Geophysics,* 56, 50–78. *doi: 10.1002/2017RG000567*

Stevens AJ, Clark D, Nicholls RJ. 2016. Trends in reported flooding in the UK: 1884–2013. *Hydrological Sciences Journal,* 61, 50–63. *doi: 10.1080/02626667.2014.950581*

UKCP Project. 2018. UKCP18 Project. *http://ukclimateprojections.metoffice.gov.uk/24125* (accessed on 10th May 2018).

University of Bristol. 2016. *Volcanic Ash Hazard Assessment for UK Nuclear Sites.* Final report, 2016-NG-D3 – Volcanic Ash. EDF Energy R&D, London, UK.

Wadey MP, Brown JM, Haigh ID, Dolphin T, Wisse P. 2015. Assessment and comparison of extreme sea levels and waves during the 2013/14 storm season in two UK coastal regions. *Natural Hazards and Earth System Science,* 15, 2209–2225.
doi: 10.5194/nhess-15-2209-2015

Wastegård S, Davies SM. 2009. An overview of distal tephrochronology in northern Europe during the last 1000 years. *Journal of Quaternary Science,* 24, 500–512.
doi: 10.1002/jqs.1269

Webb JDC, Elsom DM, Meaden GT. 2009. Severe hailstorms in Britain and Ireland, a climatological survey and hazard assessment. *Atmospheric Research,* 93, 587–606.
doi: 10.1016/j.atmosres.2008.10.034

Yan J. 2007. Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software,* 21, 1–21.

Zappa G, Shaffrey LC, Hodges KI, Sansom PG, Stephenson DB. 2013. A multimodel assessment of future projections of North Atlantic and European extratropical cyclones in the CMIP5 climate models. *Journal of Climate,* 26, 5846–5862.
doi: 10.1175/JCLI-D-12-00573.1

Akaike information criterion

Provides an estimate of the quality of statistical models for a given set of data. If several models are fitted to the same data, AIC provides a measure of the quality of each model relative to the others. AIC therefore provides a method for model selection. It is important to note that AIC does not provide a test of a model in an absolute sense — i.e. whether the fit of the model to the data is poor or not.

Copula

A multivariate distribution function, which describes the dependence between two or more variables. Copulas are based on the theory that a joint distribution of two or more variables can be described in terms of the distribution of each variable (the marginal distribution functions; see below) and a copula that describes the dependence between the variables.

Cumulative distribution function (CDF)

The CDF of a random variable U, evaluated at some value u, is the probability that $U \leq u$. Hence, the CDF has values between 0 and 1.

Kendall's tau

A rank correlation coefficient, which provides a measure of the statistical dependence of two datasets. It has values between −1 and 1. A value close to 1 would indicate that the values in the two datasets have similar or identical ranks. Values near 0 indicate little or no correlation between the datasets.

Marginal distribution function

Gives the probability of values of a random variable within a larger dataset without reference to the other variables.

North Atlantic Oscillation (NAO)

A large-scale pressure pattern that has important impacts on the weather and climate of the North Atlantic region and Europe. Its influence is strongest during winter, but can also affect summer weather. High latitudes, near Greenland and Iceland, generally experience low air pressures, whilst air pressures further south (near the Azores) are higher. The exact air pressure, and positions of the centres of the low and high pressure regions, vary over time. The NAO reflects fluctuations in the difference in pressure between these two regions. If the air pressure is lower than average over Iceland and higher than average over the Azores, the NAO is in a positive phase, which corresponds to stormy conditions and mild and wet winters over northern

Europe. If the air pressures over Iceland and the Azores are higher and lower than average respectively, the NAO is in a negative phase. A negative NAO corresponds to decreased storminess, below-average precipitation, and cold temperatures over northern Europe during winter.

Scandinavian Pattern (SP)

A pressure anomaly centred over the Scandinavian Peninsula. When the SP is in a positive phase, reduced rainfall is observed in January around the British Isles, whereas rainfall tends to be enhanced in October and April.

Return period

The period over which, on average, an event would be expected to occur. For example, an event with a 10-year return period would be expected to happen once every 10 years. Alternatively, the probability of the event occurring in any given year is 1/10.

Abbreviations

AIC	Akaike information criterion
ATDnet	Arrival Time Difference Network
CDF	Cumulative distribution function
CME	Coronal mass ejection
ECA&D	European Climate Assessment and Dataset
GPD	Generalised Pareto distribution
HadISD	Met Office Hadley Centre Integrated Surface Database
NAO	North Atlantic Oscillation
ONR	Office for Nuclear Regulation
RCP	Representative Concentration Pathway
SCP	Scandinavian Pattern
TAG	Technical Assessment Guide
TORRO	The Tornado and Storm Research Organisation
UKCP09	United Kingdom Climate Projections 2009
UKCP18	United Kingdom Climate Projections 2018

Independent events

If two (or more) events are independent, their joint probability is simply the product of their individual probabilities. As an example, the probability of transport of volcanic ash from an eruption on Iceland to the UK will be estimated. The probability depends on two independent factors: the probability of an eruption occurring on Iceland, and the probability that the winds in the North Atlantic have a north-westerly component. The latter condition is necessary, as Iceland is located at a more northerly latitude than the UK.

Sammonds et al. (2010) estimated an average frequency of eruptions on Iceland of 20 to 25 per century from historical records, corresponding to a probability of 20/100 = 0.2. They also estimated the probability of the prevailing winds having a north-westerly component as 0.06. Hence, the probability of ash being transported to the UK is 0.20 × 0.06 = 0.012. On average, therefore, transport and deposition of ash from Iceland to the UK would be expected to occur about once per century. This estimate is highly uncertain, as the amount of ash transported is also dependent on rainfall, which washes ash out of the atmosphere and reduces the amount transported. The time taken for the ash to reach the UK varies considerably with the weather conditions. Depending on the synoptic conditions, the ash could be transported northwards or southwards from Iceland before moving toward the UK. The time between the production of ash and the ash reaching the UK could be shorter than 12 hours or longer than 90 hours (*University of Bristol, 2016*).

Empirically-based methods

The average probability of a particular hazard combination could be estimated if a sufficiently long series of observations of the variables of interest is available. The number of days when the two hazards occurred simultaneously would be counted and then divided by the total number of days. For example, one hazard of interest is high winds and heavy rain (*Table 1*), where 'high winds' and 'heavy rain' will be events above predefined thresholds. If a data series consists of observations for 360 days, and 40 days were deemed to have both high winds and heavy rain, then the joint probability of high winds and heavy rain is 40/360 = 0.011. In reality, a much longer data series would be used to estimate the joint probability. This approach estimates the average probability of an event occurring, but does not give any information on the likely state at any given time. Climate change could change the frequency of some hazards. An estimation of the probability of hazard combinations based on data for the current climate may not be valid for the future climate.

If the probabilities of changing from one state to another are known (i.e. between days with and without hazards), the probability of a joint hazard occurring on the next day, or the third day ahead, can be calculated using a Markov chain. A Markov chain is a stochastic process; that is, a mathematical model that evolves over time in a probabilistic manner.

Markov chains

A simple method for estimating the long-term average probability of an event occurring was described in the previous paragraph. Here, a simple example of a Markov chain is presented, with just two states, 'hazard present' and 'no hazard'. If a hazard occurs on a given day, the following day may also have the same hazard, or be hazard-free. Similarly, if no hazard occurs on a day, the next day could be either hazard-free or contain a hazard. For this example, wet and dry days are used to represent days with and without hazards respectively. Probabilities of changing between wet and dry days, or remaining in the same state (i.e. if a day is wet, the next day is wet), are shown in *Table A1*. These probabilities were created artificially and do not represent any particular location. They could be estimated from a long time series of observed or modelled data. For joint hazards, a Markov chain with four states would be required — both hazards, either hazard, or no hazards.

Table A1. Probabilities of remaining in the same state or changing to another state.

		State on the next day	
Current State	State	Wet	Dry
	Wet	0.70	0.30
	Dry	0.20	0.80

The symbol p_{ij} represents the probability of changing from state i to state j. From *Table A1*, the probability of changing from a dry day to a wet day, p_{21}, is 0.20. Similarly, the probability of a dry day followed by another dry day, p_{22}, is 0.80. The probabilities of remaining in a given state or changing into another are often represented as a transition matrix. If P is the transition matrix then, from *Table A1*:

$$P = \begin{bmatrix} 0.70 & 0.30 \\ 0.20 & 0.80 \end{bmatrix}$$

This matrix has several properties. It is square, since all states must be represented in the rows and columns. All the entries in the matrix are probabilities, and so must have values between 0 and 1. The sum of each row must equal 1, as the numbers in the row are the probabilities of changing from a state on the left to one of the other states indicated across the top (as illustrated in *Table A1*). The transition matrix P illustrates two important points regarding Markov chains. The outcome of each experiment is one of a set of discrete states, and in the examples used here, the outcome only depends on the current state, and not on any previous states. Higher order Markov chains can be created, where the outcome depends not only on the current state, but on some number of previous states as well. For systems with a greater number of states, some of the entries in the matrix P may be zero, showing that either remaining in the same state or transitions between certain states is not possible.

The Markov chain can be used to calculate the probability of a given state several days ahead. For example, the probability of a wet day two days ahead is calculated by multiplying P by itself (P^2), i.e.:

$$P^2 = \begin{bmatrix} 0.70 & 0.30 \\ 0.20 & 0.80 \end{bmatrix} \begin{bmatrix} 0.70 & 0.30 \\ 0.20 & 0.80 \end{bmatrix} = \begin{bmatrix} 0.55 & 0.45 \\ 0.30 & 0.70 \end{bmatrix}$$

The probability of a wet day two days ahead is $p_{11} = 0.55$. The Markov chain can also be used to calculate changes in the distribution of states. For example, suppose the current distribution (initial state) of wet and dry days is 85% and 15% respectively. After one day, the new distribution can be found by multiplying the current distribution by the transition matrix:

$$\begin{bmatrix} 0.85 & 0.15 \end{bmatrix} \begin{bmatrix} 0.70 & 0.30 \\ 0.20 & 0.80 \end{bmatrix} = \begin{bmatrix} 0.625 & 0.375 \end{bmatrix}$$

The long-term distribution for day n is found by multiplying the initial state by P^n. After a number of iterations, the distribution will not change; the same long-term distribution is obtained regardless of the initial state. For the example shown here, using a time 10 days ahead (i.e. P^{10}), the final distribution is:

$$\begin{bmatrix} 0.85 & 0.15 \end{bmatrix} P^{10} = \begin{bmatrix} 0.85 & 0.15 \end{bmatrix} \begin{bmatrix} 0.40 & 0.60 \\ 0.40 & 0.60 \end{bmatrix} = \begin{bmatrix} 0.40 & 0.60 \end{bmatrix}$$

In this example, the long-term average is 40% of days to be wet and 60% dry. The same long-term average (or climatology) is obtained regardless of the initial state.

In summary, Markov chains may be used to calculate the probability of a state on the next day, or several days ahead, given an initial state. One of the states could be the combination of two hazards. In this instance, the matrix *P* would have four rows and four columns, and one row would describe the probability of remaining in a state with combined hazards (heavy rain and high winds), or moving to a state with just one, or none of the hazards. The change in distribution of states can be calculated from a given initial distribution. The long-term average distribution of states, however, is independent of the initial distribution. This last property is important, as it means that Markov chains can only be used to predict the future states for a short period; the exact length will depend on the particular hazards under study and their probabilities, and whether the next state depends on just the current state, or several previous states. In the example used here, all predictions are close to the climatology (i.e. the long-term average) after about six days.

Statistical methods

A number of more complex statistical methods exist for calculating joint probabilities of hazards, which are described briefly below.

Multivariate Bayesian framework

Sexton et al. (2011) describe a method for calculating the joint probability distribution of changes in pairs of meteorological variables based on a multivariate Bayesian framework. This method formed part of the process of creating the United Kingdom Climate Projections 2009 (UKCP09; *Murphy et al., 2009*). A complete description of this approach is beyond the scope of this technical volume; however, full details are given by *Sexton et al. (2011)*, and a summary is provided in Chapter 3 of *Murphy et al. (2009)*.

The multivariate Bayesian method combines information from different sets of climate model simulations and observations. The climate model simulations were used to train a multivariate emulator. An emulator is a statistical model that can simulate a wide range of climate variables given certain initial conditions or other information. The emulator was required because it was not possible to execute a climate model many thousands of times owing to the high computational cost. Next, 25,000 emulated simulations were selected and in turn used to set up a simple climate model, which was integrated from 1860 to 2100. The simple climate model simulations were further sub-sampled to produce 10,000 equally likely estimates of the

climate variables of interest, such as temperature and rainfall. The joint probability of changes in two variables was also calculated.

This approach for calculating the joint probabilities is computationally demanding. It involves the execution of a global climate model several hundred times, and a simple climate model many thousands of times. Observations of a wide range of climate variables are also needed to constrain and evaluate the climate simulations.

Joint tail models

Two statistical models for estimating joint probabilities of extremes have been created by *Ledford and Tawn (1997)*, and *Heffernan and Tawn (2004)*. These models offer a flexible method for capturing the range of dependence between extremes of variables. One example would be extreme swell and extreme wave heights; both of these hazards are produced by high winds blowing over the oceans. The model of *Heffernan and Tawn (2004)* has been used to estimate the dependence between extreme sea level heights at nearby locations (*Eastoe et al., 2013*).

The conditional extremes models are fitted to the data in a very similar way to the copulas (*Section 4.3*):

- Transform data so all values lie in the range 0 to 1 using a GPD above the marginal threshold and an empirical CDF below the threshold.
- Transform the data onto Laplace margins to make the dependence model fit easier (*Keef et al., 2013*).
- Set the dependence threshold, which is often taken to be the same as the marginal threshold, but doesn't need to be the same. Use the conditional extremes model to provide an estimate of the distribution of one variable (X) given that the other (Y) exceeds an extreme threshold (u); this distribution is commonly written as: $Pr(Y \mid X > u)$.
- Use the fitted model to extrapolate to higher critical levels.

These models make no assumption on the distribution of the conditional variable, i.e. if one variable is extreme, no assumption is made of the variation of the other variable.

The choice of thresholds (i.e. the thresholds for each variable, for which values larger than the thresholds are classed as extreme) are defined by the user, and should be chosen with care. The model described by *Heffernan and Tawn (2004)* requires considerable statistical expertise to

use and interpret. It has been implemented in the R programming language (see *Appendix B*).

Kernel density estimation

Kernel density estimation is a method for estimation of probability density functions. Kernel density estimators can be applied to datasets irrespective of the nature of the dependence between the variables (*Santhrosh and Srinivas, 2013*). There is a wide range of kernels that could be used. However, many kernels are best used with unbounded data, and so are less suitable for use with meteorological data that are often bounded. For example, rainfall amounts cannot have negative values, and relative humidity must have a value between 0% and 100%.

Santhrosh and Srinivas (2013) used a diffusion process-based adaptive kernel (D-kernel) to model joint distributions of variables that characterised a flooding event (peak flow, volume and duration). They provide a full mathematical description of the D-kernel and its calculation. Part of the process of construction of the D-kernel requires binning the data. *Santhrosh and Srinivas (2013)* noted that selection of the optimum number of bins is critical when using this method.

Multivariate kernel density estimation is often restricted to two dimensions (i.e. two variables), owing to a phenomenon called the curse of dimensionality. As the number of dimensions (here, meteorological variables) increases, the data space will be very sparsely populated by data points. In other words, for a given data point, there will be very few or possibly no neighbouring data points, unless the sample size is extremely large. Consequently, the error in the fitted kernel with large numbers of dimensions and very high numbers of data points is often larger than the error for a small number of dimensions and far fewer data points.

Copulas

A copula is a function that models and captures the dependence between two variables independently of their distributions. Let A and B represent sets of two variables, such as temperature and rainfall, and the total number of data values in each set is N. Their CDFs are F_A and F_B. For arbitrary limits a and b, F_A and F_B represent the probability that $A < a$ and $B < b$ respectively. These CDFs can be estimated by ranking the data and dividing the ranks by N+1 (the empirical CDF), or fitting a marginal distribution function to the data.

The joint probability of $A < a$ and $B < b$ is represented as $Pr(A < a, B < b)$. The copula, C, can be used to join or 'couple' the two individual CDFs to obtain their joint probability distribution function, $F(a,b)$:

$$F(a,b) = Pr(A < a, B < b) = C[F_A(a), F_B(b)]$$

Software packages in the R language were used to fit and evaluate the copulas (*Table B1*). All of the packages are freely available from the Comprehensive R Archive Network (*https://cran.r-project.org/index.html*). The 'copula' package is described in more detail in the paper by *Yan (2007)*. Functions in the package 'fitdistrplus' were used to fit distributions to the datasets and so create the marginal distribution functions needed for the copulas. This package also calculates AIC and other useful goodness-of-fit statistics so the quality of the fit can be judged. The 'texmex' package contains functions to fit the Heffernan-Tawn model to data. Functions to fit and manipulate copulas are also available in other languages and programming environments (e.g. Python, MATLAB).

Table B1. R software packages used to fit and evaluate the copulas.

Package name	Description
copula	Fitting copulas to datasets; evaluating copulas.
lcopula	Contains function 'K.plot' to produce a Kendall plot (or K-plot).
fitdistrplus	Routines to fit functions to data, including Weibull, gamma and normal.
texmex	Statistical modelling of extreme values, including Heffernan-Tawn model for joint extremes in two or more variables.

LC 0064_18V12

EDF ENERGY

energy technologies institute

IChemE ADVANCING CHEMICAL ENGINEERING WORLDWIDE

Institution of MECHANICAL ENGINEERS

Met Office

MOTT MACDONALD