Development and Application of a Data Synthesizer for Mobile Search of Radioactive Material Out of Regulatory Control

by

Pim Nelissen

Supervised by Christopher Rääf
Co-supervised by Christer Ceberg, Aliaksandr Dvornik and Robert Finck

Institute of Translational Medicine
Medical Radiation Physics
Lund University, Malmö, Sweden

Spring 2026

For the wellspring of life is with you;
By means of your light we see light.

Psalm 36:9

Abstract

The loss of regulatory control over radioactive material and sources poses great risks to public safety. Such a situation necessitates the development of protocols and methods for effective and accurate localisation, characterisation and securement of the materials. Mobile gamma spectrometry, particularly using vehicle-borne detectors, has emerged as a viable method for *in-situ* characterisation of the gamma radiation field, and has been successfully used in searches for locating lost gamma sources. Traditional localisation methods rely on the peak count rate and full width at half maximum of the characteristic count rate curve as a vehicle-borne detector passes a source. However, assumptions required for these methods often fail due to real-world complexities. Bayesian search methods have been developed as a robust alternative, but further validation is needed to test its effectiveness and find the limiting cases.

This thesis presents PG-RAD (Primary Gamma RADiation simulator), a modular toolbox for simulating mobile search scenarios of gamma sources. PG-RAD generates synthetic counts that would be observed by a detector travelling along a path as it passes a gamma source. PG-RAD accounts for angular detector efficiency, air attenuation and counting statistics. PG-RAD is flexible in the definition of point sources and trajectory geometry. To demonstrate its applicability to research in source localisation methods, simulations of both HPGe and NaI(Tl) detectors were performed, simulating mobile search of Caesium-137 sources. PG-RAD was validated against existing reference codes and experimental data. It was also used to systematically test a two-source Bayesian localisation algorithm, assessing its performance as a function of source separation and difference in activity.

Results show that PG-RAD is able to generate representative count data, and can effectively be used for testing mobile search algorithms. The lack of shielding and Compton scattered photons likely account for the discrepancies found when comparing to some experimental data. However, because PG-RAD is developed as an open source toolbox, taking a modular and object-oriented approach, future implementation of these features is possible to enhance the realism of the generated data. Furthermore, other detectors or isotopes could be added in order to make PG-RAD useful for other research teams around the world.

Populärvetenskaplig sammanfattning

Radioaktiva ämnen har många användningsområden i sjukvård, forskning, industri och kärnkraft. De kan dock vara farliga om materialet hanteras på fel sätt. Därför finns starka reglering av användningen över hela världen. Trots detta kommer radioaktiva ämnen ibland bort, på grund av stöld eller felaktig hantering. Radioaktivt material kommer bort oftare än vad man kanske tror. Internationella atomenergiorganet (*International Atomic Energy Agency*) tåg år 2025 emot fler än 200 anmälningar om stulna radioaktiva ämnen. I sådan fall är det viktigt att snabbt lokalisera materialet, så att man kan undvika att människor kommer till skada. En användbar metod för att lokalisera förlorade radioaktiva källor är mobil sökning. Genom att montera strålningsdetektorer på ett fordon, till exempel en bil, en helikopter eller en drönare, kan man mäta gammastrålning från radioaktiva ämnen över ett mycket stort område och snabbt söka efter strålkällor.

Syftet med detta examensarbete är att simulera scenarier för mobil sökning av radioaktiva källor. En sådan simulator är användbar eftersom forskare kan testa många olika scenarier med den, relativt snabbt och kostnadseffektivt. Simulatorn som utvecklats i examensarbetet kallas PG-RAD. I arbetet genererades ett antal olika scenarier som jämfördes med data från verkliga händelser. För att visa på möjliga framtida användningsområden användes PG-RAD även för att simulera hur en drönare skulle kunna upptäcka en gammastrålkälla från luften. Slutligen användes PG-RAD för att testa bayesianska metoder vid mobil sökning efter radioaktivt material. Bayesiansk statistik är en avancerad typ av statistik, och nya studier visar att den kan vara användbar i scenarier med mobil sökning. PG-RAD kan mycket snabbt generera många olika scenarier och framgångsrikt överföra dessa data till den bayesianska algoritmen.

Sammantaget är PG-RAD ett användbart verktyg för simuleringar av mobil sökning efter gammastrålkällor. Eftersom simulatorn är modulär är det enkelt att lägga till nya moduler eller förbättra befintliga. Framtida arbete skulle kunna inriktas på att göra simuleringarna mer realistiska genom att förbättra fysikmodellerna, som beskriver hur objekt i miljön påverkar uppfattarstrålningens spridning och absorption, till exempel i träd och byggnader.

Popular abstract

There are many use cases for radioactive materials in the modern world, such as medical care, nuclear power and industrial applications. However, radioactive sources can be very dangerous when handled incorrectly. That is why there are strong regulations all around the world to handle radioactive material securely. Despite that, radioactive material sometimes gets stolen or simply lost. It happens more than one may think: in 2025 alone, the International Atomic Energy Agency received more than 200 reports of stolen radioactive material. When a radioactive source is lost out of the regulators' control, effective localization of the material is very important to avoid harm to people. One method that has proven effective for searching for radioactive sources is mobile gamma spectrometry. By putting detectors on a vehicle such as a car, plane or drone, a much larger area can be searched than is possible with handheld detectors.

The objective of this thesis work was to develop a simulator program which can generate data that mimics mobile search scenarios of radioactive sources. Such a simulator is useful, because it allows researchers to test many different scenarios relatively quickly and cheaply. The program developed in this thesis is called PG-RAD. In the thesis, a few different scenarios were generated and compared to real-world data. To showcase the potential future uses, PG-RAD was also used to simulate how a drone might detect a radiation source from above. Finally, PG-RAD was used to test Bayesian methods in mobile search of radioactive material. Bayesian statistics is an advanced kind of statistics, and recent studies show it could be useful in mobile search scenarios. PG-RAD was able to generate lots of different scenarios very quickly, and pass this data to the Bayesian algorithm with success.

In conclusion, PG-RAD is useful for simulations of mobile search scenarios. The fact that it is a modular program means adding new modules or improving old ones is straightforward. Future work could include making the simulations more realistic by improving the physics models, and allowing for objects to be placed in the world that affect the detection of a radiation source, such as vegetation and buildings.

Contents

1 Introduction

It is well understood that high doses of ionizing radiation can have adverse biological effects on the human body. Therefore, proper handling of radioactive material, along with proper education, training and security measures, are essential for ensuring the safe application of nuclear technology in society. Because of the possible risks involved, the nuclear technology sector is highly regulated by international and national agencies alike. The International Atomic Energy Agency (IAEA) is one such agency, monitoring and regulating the nuclear sector in over 180 member states.

Despite extensive regulation, various incidents involving the release of radioactive material into the environment have occurred in the past, some of which have resulted in injuries, illness and casualties. The IAEA uses the term Material Outside of Regulatory Control (MORC) for such a scenario [1]. In 2025, the IAEA received over 200 reports of incidents involving MORC, with most of these reports involving theft of radioactive sources from industrial or medical environments [2]. The IAEA is also involved in monitoring situations in war zones, because damage to civilian and industrial infrastructure such as hospitals, mining or nuclear power plants introduces a serious risk for highly radioactive material to get out of regulatory control.

There are different types of ionising radiation, of which gamma radiation is of particular interest. Gamma radiation is emitted by many types of radioactive sources that are used in industries and hospitals. Gamma photons can travel relatively long (hundreds of meters) through air before depositing their energy and therefore pose a risk for tissue damage in humans. The very first observations of gamma radiation date to the late 19th century, when Wilhelm Conrad Röntgen first observed the effect this type of radiation had on photographic film [3]. The field of *gamma spectrometry* is a category of analytical techniques that revolves around identifying and quantifying radioactive materials by measuring the energy and photon fluence rate (photons per unit of area and time) of gamma radiation. This field started developing in the 1950s [4], enabled by developments in detector material science and digital data processing. The first outdoor measurements of the gamma field using NaI(Tl) detectors were performed in the 1960s, and high-resolution Germanium-based detectors first started being used for field measurements in the 1970s [5].

However, characterizing and locating MORC in the field presents unique challenges that are not found in the lab. *In-situ* (on site) gamma spectrometry has to account for gamma interactions with matter such as vegetation, buildings, and the air itself. The presence of these objects complicates measurement and subsequent interpretation of results. Another aspect is the practicality of setting up detectors in the field, as some types of detectors are sensitive to changes in humidity, temperature and vibrations [6]. When there is a need to map the gamma radiation in a larger area, or there is risk for endured exposure of personnel, *mobile gamma spectrometry* may be employed. Various research teams around the world are actively developing mobile search methods, which involve the use of gamma radiation detectors mounted to vehicles such as cars, helicopters or drones.

1.1 Historical radiological incidents

To highlight the relevance of mobile search methods, let us consider these historical examples of radioactive MORC.

Kosmos 954 satellite re-entry (1978) On January 24 1978, the nuclear-powered Soviet satellite Kosmos 954 crashed in Northwest Territories of Canada. An operation known as Operation Morning Light was started in collaboration with U.S. teams to recover the debris. In the first phase of the search, survey flights were performed in a grid pattern over a search area of over 75000 km^2. Four C-130 Hercules aircraft equipped with sodium iodide based detectors were used for this initial grid search [7]. This initial survey narrowed down the search area. Low flying helicopter aircraft were used to survey this smaller area, with helicopter crew noting the location of elevated radioactivity levels. After several months of clean up effort, about 65 kilograms of material were found, although the satellite likely weighed several tonnes. A few of the recovered fragments were highly radioactive and could have caused harm or death to people under prolonged exposure, although no injuries or deaths have been linked to the fallout from the Kosmos 954 crash.

Chernobyl disaster (1986) On April 26, 1986, tests were being performed on reactor 4 of the Chernobyl Nuclear Power Plant in Northern Ukraine. To perform the test, the automated shutdown system was disabled, which violated protocol. Control rod operations which were performed during this test caused large fluctuations in temperature and flow rate of the incoming cooling water. In particular, a design choice where control rods initially increase reactivity upon insertion, combined with the ^{135}Xe poisoned state, led to a power surge once operators withdrew the control rods. This power surge was far beyond design specifications. The operators were not permitted to perform this operation under these conditions, but due to lack of training the operators were not aware of this. and prompt criticality of the reactor caused a runaway reaction resulting in graphite fire and damage to the core vessel, exposing the core to the atmosphere. This caused a difficult-to-extinguish fire, releasing radionuclides into the environment for several days before it could be extinguished. A large variety of radionuclides were released into the environment, but from a public health perspective, the two most important ones were the short-lived isotope ^{131}I and the longer-lived ^{137}Cs [8]. The former caused significant increase of thyroidal cancer cases in the vicinity. As for ^{137}Cs, anomalous levels were found as far north as the arctic regions of Finland and Sweden, where elevated levels were found in reindeer milk [9].

Goiânia teletherapy source theft (1987) A ^{137}Cs source with an activity of 50.9 TBq was stolen from a teletherapy unit located in an abandoned hospital in Goiânia, Brazil [10]. Identifying the old hospital equipment as potentially valuable, unauthorized people entered the abandoned building and took the equipment containing the source. The ^{137}Cs source was sold to a junk yard, where it was subsequently removed from its shielding. The owner believed the caesium chloride salt, with its faint blue-ish glow, could be of value. It was also believed by some that there were mystic qualities to the object. Small pieces of the source were distributed to various people as gifts. The source ended up being exposed for over two weeks. The incident was discovered when the mother of the worst affected family brought the capsule containing the main part of the radioactive caesium chloride powder to a health centre, stating that it was killing her family. A physicist, alerted by a doctor, confirmed the radioactivity of the material using a prospecting instrument. This triggered a large-scale search for dispersed ^{137}Cs that had leaked from the source capsule since the theft. Over 100000 people were checked for exposure, of which 249 people had been found to have been contaminated. Over 100000 people were checked for exposure, of which 249 people had been found be to have been contaminated. Elevated levels of ^{137}Cs were found in sediments of the nearby river, and contamination was identified in residences up to 100 km away. The aftermath of this incident tragically resulted in four deaths due to exposure to the source.

Western Australia gauging source incident (2023) In January 2023, an 8 mm by 6 mm radioactive capsule, containing a 20 GBq ^{137}Cs source used for gauging the density of iron ore, was lost along a 1400 km journey from the Rio Tinto Gudai-Darri iron mine to Perth [11]. A few days after arrival in Perth, when the inventory of the transported goods was assessed, the source was unable to be accounted for. Once the missing source was reported, response teams started the search along the stretch of highway between the mine and Perth. It took a total of two weeks to find the missing source, which posed a serious risk to civilians due to the strength of the source.

1.2 The use of *in-situ* and mobile gamma spectrometry

In the cases described above, some form of either *in-situ* (on-site) or (early form of) mobile gamma spectrometry was used. The survey teams in Northern Territories of Canada used airborne gamma spectrometry methods to locate the debris of the Kosmos 954 satellite. In Sweden, the measurement of Chernobyl related fallout was performed by *In-situ* measurements using high-purity germanium detectors and an airborne NaI(Tl) detector was used for larger area surveys [4, 12]. In the Goiânia incident,*in-situ* measurements were performed across a large area of the province, mapping contaminated areas and people by detecting elevated radiation levels using hand-held detector instruments [10]. For these three cases, GPS navigation technology was not available yet, and thus mobile gamma spectrometry using cars in the way it is performed today was not possible. By the time of the 2023 Australia incident however, live mobile gamma spectrometry, which uses GPS information, was already a proven method. A gamma detector was mounted inside a car, and the team collected measurements at a velocity of 70 km h^{-1} across the 1400 km section of highway. The staff in the vehicle identified elevated gamma radiation at a particular section of highway, motivating a search with a handheld detector in that area in order to find its exact location [11].

1.3 Development of Bayesian methods in mobile gamma spectrometry

In the radiological emergencies mentioned previously, vast areas had to be searched. Such scenarios benefit from mobile setups, which allow searching a broad area for gamma radiation. Furthermore, mobile setups reduce the time needed for staff to set up and record data at stationary *in-situ* points. One such solution is mounting gamma detectors to a car. Such a set up introduces new real-world effects, such as the effects of velocity, reduced counting statistics, distortions, shielding effects from buildings and viaducts, and variation in the environmental radiation as the vehicle moves through a landscape.

Although it is possible to describe the expected signal from a mobile gamma spectrometry setup analytically, assumptions and simplifications that are required for this are often violated by the aforementioned real-world effects. Various research groups are therefore investigating the use of Bayesian methods in gamma spectrometry. Hite and Mattingly [13] and Kim et al. [14] have applied Bayesian statistics to infer positional information of radioactive sources using distributed networks of detectors.

Bukartas et al. [15] have shown that it is technically feasible to use Bayesian statistics for inferring the location and activity of MORC in the form of an orphaned point source using a mobile detector setup. Recent work by Dvornik et al. [16] has highlighted how Bayesian statistics can be advantageous in cases that are typical for mobile search scenarios of orphaned

point sources, such as those involving large search areas and a low signal-to-noise ratio. Although the Bayesian methods have been successfully applied in mobile gamma spectrometry, work is still ongoing to refine the method and potentially make it deployable in emergency scenarios in the future.

1.4 Objectives of the thesis work

Developing an orphaned source simulator

One of the most efficient ways to validate and improve a model in the early stages is to use synthetic data. The primary goal of this thesis work is to develop a new simulator program for mobile gamma spectrometry scenarios of orphaned sources that can

- simulate path and source geometries in a 3D Cartesian landscape,
- numerically approximate the data acquired when a detector travels along the specified path, and
- generate a wide variety of such scenarios in order to support validation and testing of the Bayesian methods under development.

In order to be useful beyond the work of this thesis, the software should be built in such a way that it is maintainable and extensible in the future. This means there are a number of requirements that should be satisfied:

- The code should be well written, in a programming language that is widely known, and written in a standard way.
- The synthesized data should be representative of real-life. This means factors like detector efficiency, air attenuation of gamma rays, and counting statistics need to be accounted for.
- The simulator should be usable by programmers and non-programmers alike. Scenario creation should therefore be flexible while avoiding requirement for scripting.

Application of the simulator

The second part of the thesis work is to show use cases of the simulator, particularly using the generated data to explore the Bayesian methods for orphan source localization as described by Bukartas et al. [15] and Dvornik et al. [16]. To build up to that, data generated by the simulator will be compared with other data synthesizers. Next, experimental scenarios will be recreated in order to see whether the simulator can generate data resembling real-life scenarios as well. Finally, various use cases of the simulator are shown, including using the simulator to test Bayesian methods for source localisation.

2 Physics background

Atoms are composed of a positively charged nucleus and a negatively charged 'cloud' of electrons surrounding it. The nucleus itself is composed of protons (symbol Z) and neutrons (symbol N). Z is the atomic number that determines which *element* an atom is composed of. Nuclei can be symbolically written in the form ${}^{A}_{Z}X$, where X is the symbol that is based on the name, and $A = N + Z$. As an example, ${}^{12}_{6}\mathrm{C}$ denotes the element carbon, which has 6 protons and $A = 6 + 6 = 12$ nucleons.

Atoms that have a different number of neutrons N but the same number of protons Z are called *isotopes* of that element. For example, the element caesium ($Z = 55$) has various isotopes including ${}^{133}\mathrm{Cs}$, ${}^{134}\mathrm{Cs}$, ${}^{135}\mathrm{Cs}$ and ${}^{137}\mathrm{Cs}$. When referring to a specific isotope, including its configuration of nucleons as well as the electrons, one speaks of the *nuclide*.

Most nuclides that are found in nature are stable, meaning the configuration of nucleons in the nucleus does not change over time. However, some nuclides are unstable due to the nucleus having excess energy, which is called being in an *excited state*. It is then energetically favourable for the nucleus to get rid of this energy, reducing its excitation and reach a more stable configuration. The emission of one or more particles to reduce excess energy is known as *radioactive decay*, and nuclides that can undergo such a decay are called *radionuclides*.

2.1 Gamma radiation

Radionuclides can emit various types of radiation, but the one most relevant for mobile searching of radioactive material is gamma radiation. A gamma decay can be described as

$$ {}^{A}_{Z}X^{**} \rightarrow {}^{A}_{Z}X^{*} + \gamma. $$

Note that it is typically assumed that the energy loss due to recoil of the emitting nucleus is negligible, meaning that the energy of the emitted photons is equal to the difference between the two nuclear states ${}^{A}_{Z}X^{**}$ and ${}^{A}_{Z}X^{*} + \gamma$. Emitted gamma ray(s) from radionuclides are thus specific to the particular excited nucleus and its internal nuclear structure.

Unlike other types of radiation, gamma rays are detectable over large distances and gamma emission is observed in all nuclei that have excited states ($A > 5$) [3, 17, 18]. Because the emitted gamma rays are characteristic to the radionuclide, they can serve as a sort of 'fingerprint' for the nucleus. For example, ${}^{137}\mathrm{Cs}$ and ${}^{60}\mathrm{Co}$ are two gamma emitting radionuclides. ${}^{137}\mathrm{Cs}$ emits a 661.7 keV gamma ray, and ${}^{60}\mathrm{Co}$ can emit a 1173 keV and a 1332 keV gamma ray. Detecting anomalous amounts of, say, 661.7 keV gamma rays in the field, can thus indicate the presence of a ${}^{137}\mathrm{Cs}$ source nearby.

2.2 Natural sources of radionuclides and gamma emitters

Each person on Earth is exposed to natural radiation. The two primary sources of natural radiation are discussed below.

Cosmic rays High energy particles from outer space, called *cosmic rays*, reach Earth from distant galaxies. Arriving at Earth, these cosmic rays can interact with the particles in the at-

mosphere and the Earth's surface. These interactions can produce a wide variety of radionuclides, some with half-lives of only a few minutes, but others that live for longer. Examples of such isotopes are ^{3}H, ^{7}Be, ^{14}C, and ^{22}Na [19]. This form of radiation is particularly relevant for gamma spectrometry, as some of these radionuclides (particularly ^{7}Be and ^{22}Na) produce gamma radiation. However, due to their low natural abundance, there is no noteworthy contribution to the background in either *i situ* or mobile gamma spectrometry.

Primordials From the Latin *prīmōrdiālis*, meaning 'from the beginning', *primordial radionuclides* have been part of the material constituents of the Earth before it was formed, or are a decay product resulting from such constituents. The latter occurs when radioactive primordials decay into another radioactive nuclide. These daughter products also fall under naturally occurring radionuclides. Some examples of primordial radionuclides include ^{40}K and ^{87}Rb. ^{232}Th and ^{238}U, with their corresponding daughter products also considered primordials [4]. Of these, the primary contributors to radiation exposure to the public are ^{40}K and ^{222}Ra, the latter of which is a daughter product from the ^{238}U decay chain [19, 20]. These primordials have significant effect on measurements performed with mobile gamma spectrometry, particularly for lower energy resolution detectors [4].

On average, the cosmic ray products and the primordials account for a dose of about 2.4 mSv/ year, which is more than 80 % of the yearly dose received by the public [20].

2.3 Anthropogenic sources of radionuclides in the environment

The synthesis of radionuclides by humans is done for a broad range of activities. Such activities result in the release of radionuclides into the environment in various ways, both in intentional and unintentional ways. A few of the main activities of interest are mentioned here.

Nuclear weapons testing The atmospheric nuclear weapons tests performed between the 1940s and the late 1970s have contributed significantly to certain populations of radionuclides on Earth. The fission processes in these explosions released significant amounts of certain long-lived isotopes such as ^{3}H, ^{90}Sr, ^{14}C and the gamma emitter ^{137}Cs into the atmosphere [19]. The latter is still present in the environment in detectable amounts in soil and organic materials, thus contributing to background radiation in mobile gamma spectrometry scenarios.

Nuclear power plants As of late 2025 there were 416 operational nuclear power plants (NPPs) around the world, with a total capacity of 376.3 GW(e) [21]. With the recent renewed interest and investment [22] in nuclear power around the world, it will continue to be an important source of anthropogenic radionuclides in the environment. During normal operation of an NPP, the fission process produces by-products, many of which have long half-lives. Certain neutron activation products, such as ^{3}H and trace amounts ^{60}Co get released into the environment during operations. Unless there is any damage to the fuel and/or the vessel, the most active and dangerous radionuclides remain within the fuel assembly and do not leak out to the environment during normal operations. Most of the NPP-related radionuclides in the environment are released at other stages in the fuel cycle. This includes the stages of mining, enrichment, fuel reprocessing and eventual geological disposal [23].

Spent fuel reprocessing Spent fuel reprocessing sites such as those at Sellafield, UK and Cap de La Hague, France handle spent fuel from NPPs. Through chemical processes, these facilities extract reusable materials such as uranium and plutonium, separating it from the fission by-products. Measurable amounts of fission by-products enter the environment from such facilities, mostly through marine discharges [24]. However, these marine discharges are not very relevant for mobile gamma spectrometry.

Nuclear medicine The field of nuclear medicine is vast, but can broadly be split into diagnostics and therapy [18]. An example of the first is X-ray tomography. Diagnostics is the primary contributor to public dose. Therapy contributes to a lesser extent. Although the radionuclides used in nuclear therapy can be very high activity, they typically have a short half life. For such isotopes, possible incidents or improper disposal do not form significant somatic risks [25]. Long-lived and high activity radionuclides, such as ^{137}Cs and ^{60}Co sources, are however used in blood irradiation therapy and for calibration of medical equipment [19]. As discussed in the introduction, there is a risk of theft or improper handling of such radioactive sources, and thus this is a potential source of radioactive MORC.

Industry Radionuclides are also used in a wide variety of civil industries, at various stages of their processes. The Swedish Radiation Safety Authority (*Strålsäkerhetsmyndigheten*, SSM) has registered 700 licensed users of radioactive sources in Sweden. This amounts to thousands of individual radioactive sources in use across the country [25]. Examples of such non-nuclear industries are steel, mining, chemical processing and food technology. This universal presence of radioactive sources across industrial society further highlights the need for development of (mobile) search methods for locating radioactive material that gets out of regulatory control.

Nuclear incidents and disasters In the case of fuel damage or other significant incidents at nuclear facilities, the fission products produced by the neutron induced fission of ^{235}U and ^{239}Pu in the nuclear fuel can get released into the environment. This can have major consequences for human health and the environment. The immediate response in such a situation should include mapping of the radiation, for which mobile gamma spectrometry is a useful method. Even beyond the immediate fallout, long-lived radionuclides can continue to have an effect on *in-situ* and mobile gamma spectrometry. The fallout of ^{137}Cs resulting from the 1986 Chernobyl disaster is still omnipresent in the environment of large parts of Europe, and is often detected in mobile gamma spectrometry.

2.4 Detection of gamma rays

As discussed in Section 2.1, radioactive sources can be identified by the characteristic energy of the gamma ray(s) that it emits. The actual detection of these gamma rays is done using radiation detectors. There are two classes of detectors that are commonly used in gamma spectrometry: scintillation detectors and semiconductor detectors. These detectors are made with materials such as NaI ($Z_{eff} = 53$) and Germanium ($Z = 23$), which are sensitive to gamma radiation. They do however have fundamental differences in the detection principle, making the detectors different but complimentary.

In *scintillator detectors*, the incoming gamma rays are absorbed, exciting the atoms in the detector material. These excited atoms then de-excite by emitting visible or ultraviolet light, a process referred to as scintillation. The intensity of the produced scintillation light is proportional to the deposited energy E_γ^{dep}. This light is collected by a photo-multiplier tube and then

electronically recorded. NaI(Tl) is a common material for scintillator detectors. Due to their high detection efficiency and ability to work without active cooling, NaI(Tl) detectors are a popular choice for mobile gamma spectrometry.

Semiconductor detectors rely on an entirely different process. When gamma rays interact with the semiconductor material such as high-purity germanium (HPGe), it ionizes atoms, creating electron-hole pairs. A voltage bias across the detector causes these electrons and holes to be collected, producing a pulse proportional to the energy deposited by the gamma ray. Gamma rays are thus directly converted into an electrical signal. HPGe detectors have been gaining popularity in mobile gamma spectrometry because of their superior energy resolution compared to scintillation detectors, as can be seen in Figure 1. HPGe detectors do require active cooling however.

2.4.1 Gamma ray interaction with matter

In an ideal case, the gamma rays deposit their entire energy in the detector, giving a distinct peak in the spectrum which allows one to identify the source and approximate its activity. In reality, however, full energy deposition is not the only possible outcome. Gamma rays can interact with the detector material through several different processes, resulting in other energy depositions which complicates the spectrum, as counts show up at other energy bins. The dominant interaction depends on the gamma ray energy and the atomic number Z of the detector material [3].

The photoelectric effect results in full absorption of the gamma ray and produces the full-energy peak in the spectrum. Compton scattering instead transfers only part of the gamma-ray energy to the detector, and the scattered photon may escape or scatter off another nucleus in the detector. These single or multiple scattering events lead to what is known as the Compton continuum. In mobile gamma spectrometry, Compton scattering is particularly relevant, because of the small detector volumes in mobile setups. Pair production can occur for gamma-ray energies above 1022 keV, but is not as relevant for mobile search of orphaned sources.

Due to these interactions, the measured pulse height spectrum contains both full-energy peaks and background contributions from partial energy deposition, the latter being present across the entire energy range that the detector is sensitive to.

2.4.2 Attenuation

Gamma rays do not *only* interact with the detector material. Gamma rays emitted from a source will also interact with the air and any potential surrounding materials before reaching the detector. Although interactions of individual photons are determined on the quantum level, treating the photons as a continuous flux of N_0 photons emitted from a source, one can use the Beer-Lambert law to describe the reduction of photons as a function of distance travelled through the medium

$$N(x) = N_0 e^{-\mu x}, \tag{2.1}$$

where μ is known as the *linear attenuation* coefficient of the medium. In the context of mobile gamma spectrometry, attenuation in air is moderate, but becomes more significant with larger detector-to-source distances. The linear attenuation coefficient in air, μ_{air}, can be calculated

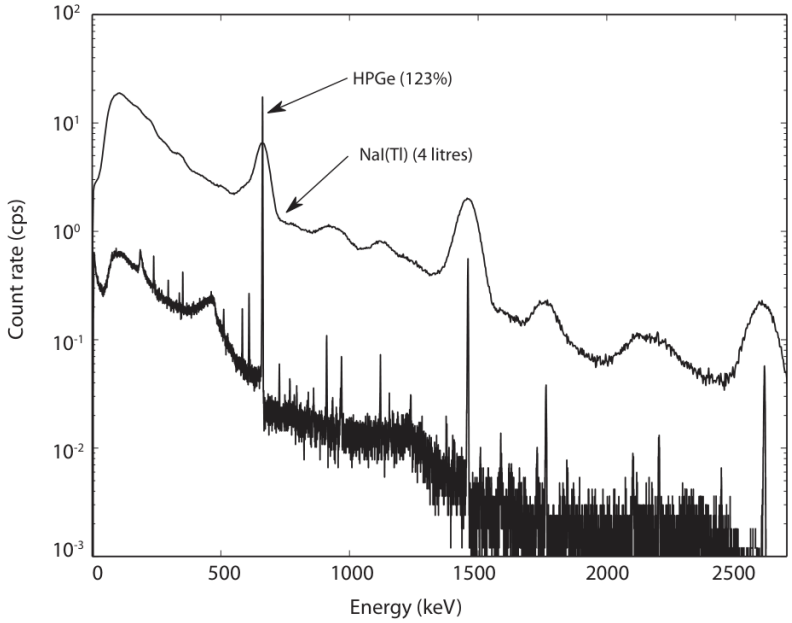


Figure 1: A pulse height spectrum recorded in front of a ^{137}Cs source. The primary gamma line is visible at around 662 keV. The HPGe detector can be seen to have a significantly better energy resolution than the NaI(Tl) detector. Figure reproduced from [19].

from the density of air ρ_{air} in the environment the measurement is taken, and the mass attenuation coefficient $\mu_{mass,air}$ as $\mu_{air} = \rho_{air} \cdot \mu_{mass,air}$ [26].

2.4.3 Recording measurements

The digital systems of a detector setup read out the deposited energy, storing it in discretized energy bins. The number of pulses within a bin as a function of the energy gives a histogram resulting in a pulse height spectrum as seen in Figure 1. In mobile gamma spectrometry, the signal of the histogram is typically not acquired continuously, but rather in set acquisition intervals on the order of 1-10 seconds. This will be discussed in more detail in Section 2.5.

2.4.4 Detector efficiency

Field efficiency The field efficiency ε_d is the intrinsic fraction of incident photons which are converted to a recorded count. The field efficiency thus is related to the effective detection area and acts as a scaling factor to the fluence rate (unit $\mathrm{m}^{-2}\,\mathrm{s}^{-1}$). Two detectors of the same model and manufacturer can have different field efficiencies. This is especially true for HPGe detectors, due to their monolithic crystal structure. Therefore, it is typical to determine the field efficiency by using calibration point sources of various primary gamma energies, placing them at a fixed distance to the detector. This creates a detector-specific field efficiency curve $\varepsilon_d(E_\gamma)$ in the unit m^2 as seen in Figure 2.

Angular efficiency Angular efficiency, typically expressed as a relative efficiency $[0, 1]$, refers to the change in efficiency as a function of the angle of incidence θ of the photon, relative to the angular positioning of the detector θ_0. This is primarily impacted by the geometry in which the detector is placed in the vehicle. Like with the field efficiency, angular efficiency is typically determined by characterising it while the detector is situated in its experimental

setup, so that attenuation from the car structure can be accounted for. An example is shown in Figure 3.

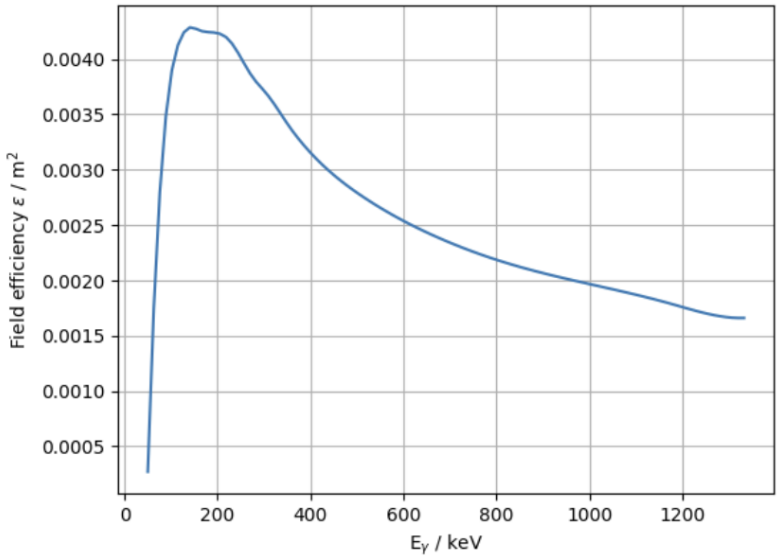


Figure 2: Field efficiency $\varepsilon_d(E_\gamma)$ of an HPGe detector mounted in a vehicle.

2.5 Locating a source using mobile gamma spectrometry

To better understand the procedure for identifying an orphaned source in the field using mobile gamma spectrometry, consider the following scenario: A single unshielded ^{137}Cs source is lost during transport, and is located a distance d away from the road. An emergency response team prepares a vehicle with a detector (say, a NaI(Tl) detector). The scenario is sketched in Figure 4.

Because of details on the freight list of the transport company, the search team know the source is a ^{137}Cs source. Therefore, a certain region of interest (ROI) of energy bins can be selected (say a few bins around the peak of 662 keV). As the detector moves along the path, the counts in the ROI $N = (N_1, N_2, , \ldots, N_k)$ are measured at locations $X = (X_1, X_2, \ldots, X_k)$. In order to obtain usable data, the team needs to make sure that a) sufficient data is accumulated per recorded point to be statistically useful; and b) the velocity of the car is not too large. There is an inherent trade-off between spatial resolution and counting statistics. While shorter acquisition times improve spatial resolution, they also increase statistical uncertainty. Longer acquisition times yield smoother spectra but lack spatial resolution needed to resolve the position of a lost source. A typical acquisition time t_{acq} in mobile gamma spectrometry is 1 to 10 seconds [28]. Let us assume the team's system records at 1 s intervals. The team drives at a velocity of 30 km h^{-1}.

Figure 5 shows what the true count rate might look like as the team passes the source. Assuming a straight road segment and a constant velocity, the shape of the true count rate curve is dependent only on the distance to the road [16, 27], while the height of the peak is proportional to the activity. Therefore, the minimum distance to the road and the activity can be estimated from the peak shape, after correcting for background radiation.

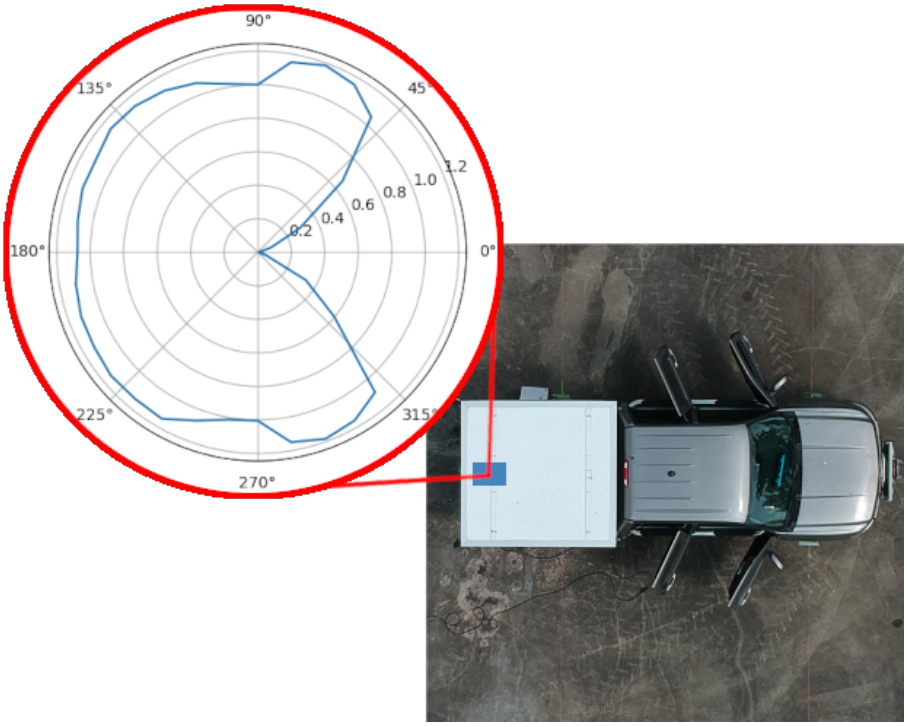


Figure 3: Relative angular efficiency $\varepsilon_\theta(\theta)$ of an HPGe detector mounted in a vehicle. The angle is relative to driving direction of the vehicle, and determined at a relative height of 1 m above the calibration source (embedded in the vehicle). The shielding caused by the vehicle's metal structure and the detector crysostat results in an apple-shaped angular efficiency curve with an efficiency of nearly 0 for photons coming in from the direction of travel. The relative efficiency is for a 661.7 keV incoming gamma photon. The photo of the car is reproduced from [27].

With discrete acquisitions however, velocity and acuisition time affect the as is done in this scenario, the relative alignment of the acquisition windows with respect to the source can affect the acquired. This is shown in Figure 6.

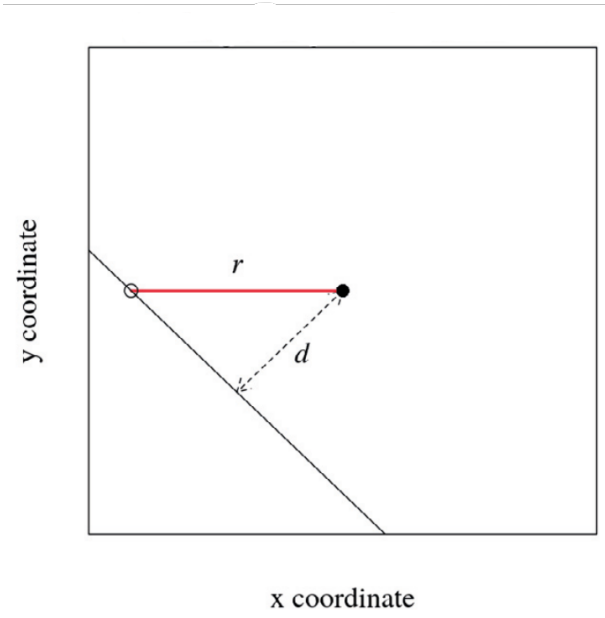


Figure 4: An unshielded point source, marked in black, is passed by a detector at a constant velocity along the red dashed line. As the distance to the source r decreases, the observed count rate increases with a peak at r_{min}. The count rate curve is shown in Figure 5. Figure reproduced from [27].

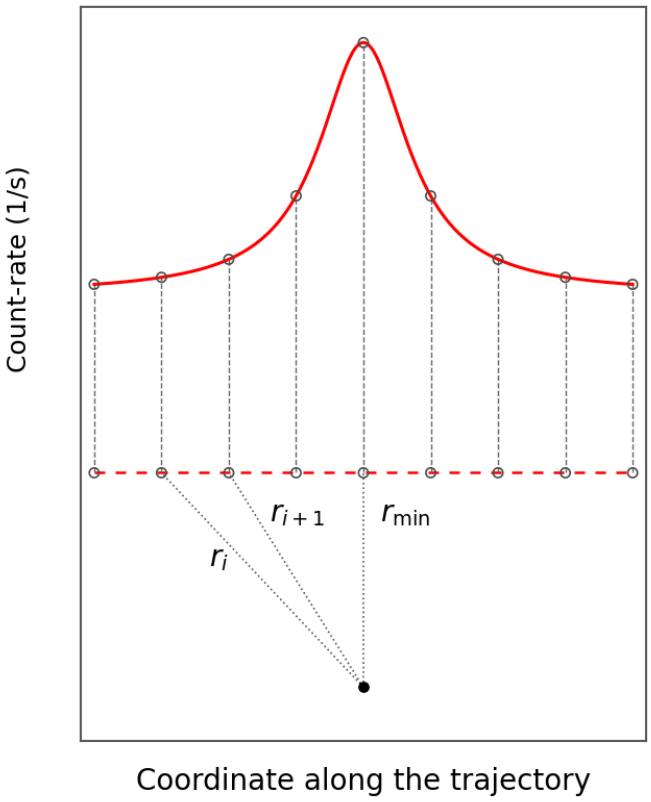


Figure 5: The expected count rate curve from a scenario like in Figure 4. Assuming a constant velocity, the expected shape of the count rate curve is a symmetric bell shape.

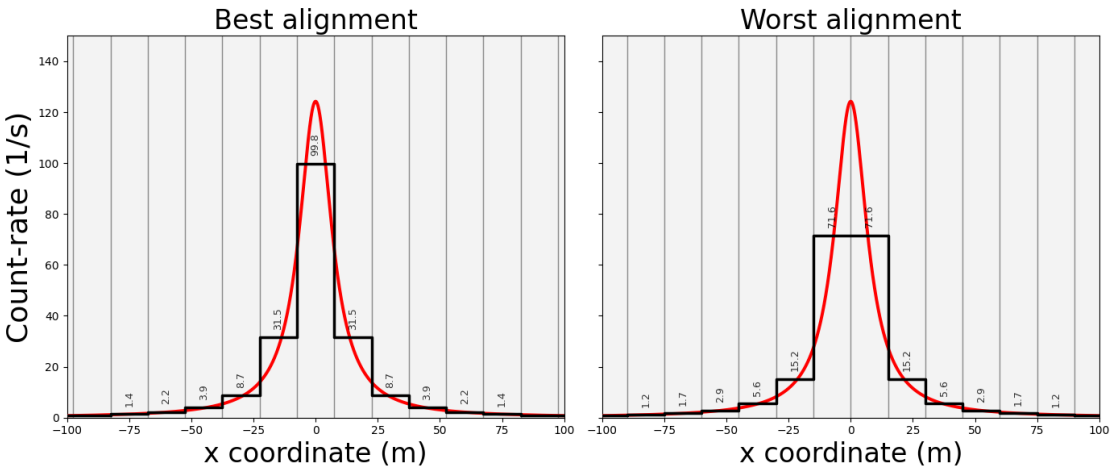


Figure 6: The best and worst possible case of alignment of acquisition windows, in the case of a trajectory perpendicular to the x-axis. The space is divided into equally sized bins. The red line shows the continuous count rate function along the trajectory. The bins show the integrated count rate that would be acquired in a finite acquisition window with acquisition time t_{acq}. One can see that (mis)alignment of the acquisition windows compared to the source location can significantly affect the recorded peak counts, and thus lead to a potential underestimation of the true activity of the source.

3 Mathematical background

3.1 Mathematics of radioactive decay

On the most fundamental level, radioactive decay is governed by the laws of quantum mechanics. It is not possible to know when an individual nucleus will decay, because this is a random process. However, radioactive sources have thousands of trillions of radioactive nuclei per cubic centimetre. When describing the behaviour of such large populations of radionuclides, it is possible to describe the expected number of decays per unit of time. The number of expected decays per second is known as the *activity A* of a source (unit Bq). This is one of the primary quantities of interest in a mobile search scenario. As such, the remainder of this section will build up the mathematics to describe the activity, that is, number of expected decays in a time interval $[t, t+dt]$.

3.1.1 Probabilistic formulation of radioactive decay

From nuclear physics principles, it is assumed that the probability for a nucleus to decay is

(A1) only dependent on its current state at time t,

(A2) independent of other nuclei and their states, and

(A3) constant over a small time interval $[t, t+dt]$.

Let T be the random variable that determines when a given nucleus decays. Then the probability that the nucleus decays in a time interval $[t, t+dt]$, given it has survived until time t, is described by

$$P(t \leq T < t+dt \mid T \geq t) = \lambda\, dt \tag{3.1}$$

where λ is the constant decay rate (A3). Notice the probability is only dependent on the constant λ and dt (A1-A2). An equivalent statement is

$$P(t \leq T < t+dt) = \lambda P(T \geq t)\, dt = \lambda S(t)\, dt \tag{3.2}$$

where $S(t)$ is called the *survival function*. Performing basic set arithmetic gives

$$P(t \leq T < t+dt) = P(T \geq t) - P(T \geq t+dt) = S(t) - S(t+dt). \tag{3.3}$$

Combining Equation 3.2 and 3.3 gives

$$S(t+dt) - S(t) = -\lambda S(t)\, dt. \tag{3.4}$$

Letting $dt \to 0$, this can be written as the differential equation

$$\frac{dS(t)}{dt} = -\lambda S(t)\,. \tag{3.5}$$

By definition, $S(0) = 1$, which means the ODE in Equation 3.5 has the well-known solution

$$S(t) = \exp(-\lambda t)\,, \tag{3.6}$$

3.1.2 Radioactive decay law

Let N_0 denote a large population of radionuclides. By (A2), the expected number of nuclei remaining after time t is $N(t) = N_0 S(t)$. Equation 3.6 can thus be equivalently written for such a population:

$$N(t) = N_0 \exp(-\lambda t)\,, \tag{3.7}$$

where $N_0 = N(0)$ is the initial population and λ signifies the decay rate. The activity (symbol A, unit Becquerel, where $1\,\mathrm{Bq} \equiv 1\,\mathrm{s}^{-1}$.) can then describe the number of expected decays per second as $A(t) = \lambda N(t)$, which means that

$$A(t) = A_0 \exp(-\lambda t) \tag{3.8}$$

where $A_0 = \lambda N_0$. The number of observed counts ΔN in a time interval $[t, t + t_{obs}]$ is then

$$\langle \Delta N \rangle = \int_t^{t+t_{obs}} A(t) dt\,. \tag{3.9}$$

If $t_{obs} \ll t_{1/2}$, then the $A(t)$ is approximately constant, meaning $\Delta N \approx A t_{obs}$. In this case, the approximation[1] can be quantified in terms of probability as

$$P(\Delta N) = \frac{\langle \Delta N \rangle^{\Delta N} \exp(-\langle \Delta N \rangle)}{\Delta N!}. \tag{3.10}$$

where $\Delta N = A t_{obs}$. Equation 3.10 is a Poisson distribution. The variance of a Poisson process is $\sigma^2 = \langle \Delta N \rangle$.

3.2 Mathematics of mobile gamma spectrometry

3.2.1 Primary photon fluence rate

Let $\vec{r}_p = (x_p, y_p, z_p)$ denote the location of an unshielded, isotropic point source p of activity A (MBq), with a known isotope with primary gamma branching ratio n_γ. Let $\vec{r}_i = (x_i, y_i, z_i)$ denote an arbitrary point in space. The primary photon fluence rate at $\vec{r}_i$ is then given by

$$\dot{\phi}(r) = \frac{A n_\gamma \exp(-\mu_{air} r)}{4\pi r^2} \tag{3.11}$$

where $r = ||\vec{r}_p - \vec{r}_i||$, μ_{air} is the linear attenuation coefficient for air, as described in Section 2.4.2. The units are $\dot{\phi} \sim \frac{\text{photons}}{s \cdot m^2}$.

[1] Radioactive decay is better described by a Binomial distribution, which can be approximated by a Poisson distribution under two conditions: (1) the population $N_0 \to \infty$; and (2) that the number of decaying particles ΔN in a time frame $[t, t + t_{obs}]$ is much smaller than the total available ($\Delta N \ll N(t)$). Furthermore, in typical gamma spectrometry scenarios, the detection efficiency ϵ is relatively small, which further reduces the discrepancy between the binomial distribution and its Poisson approximation. For these three reasons, the Poisson approximation will continue to be used, since all three of these conditions are satisfied in typical mobile gamma spectrometry scenarios. The interested reader is referred to Sitek and Celler [29] for a detailed explanation on the conditions when the Poisson approximation of counting statistics in radioactive decay fails.

3.2.2 Count rate in a detector

Detectors do not count all photons that reach it, because the electromagnetic interactions of photons with nuclei is of a stochastic nature [19]. As discussed in Section 2.4.4, this can be expressed by:

- the field efficiency for detecting primary gamma photons $\varepsilon_D(E_\gamma)$ in units of area m^2, and
- the relative angular efficiency $\varepsilon_\theta(E_\gamma, \theta) \in [0, 1]$, dimensionless.

The angular dependent absolute efficiency of the detector is then defined as

$$\varepsilon(E_\gamma, \theta) = \varepsilon_D(E_\gamma)\varepsilon_\theta(E_\gamma, \theta)\,. \tag{3.12}$$

Where $\varepsilon(E_\gamma, \theta) \sim \mathrm{m}^2$.

The *count rate* of full energy photons in the detector $\dot{N} \sim \frac{\text{counts}}{s}$ is then the fluence scaled by this efficiency $\varepsilon(E_\gamma, \theta)$. If the detector D is positioned at $\vec{r}_i$, then

$$\dot{N}(r, E_\gamma, \theta) = \varepsilon(E_\gamma, \theta)\phi(r) = \varepsilon(E_\gamma, \theta)\frac{An_\gamma \exp(-\mu_{air} r)}{4\pi r^2} \tag{3.13}$$

3.2.3 Integrating count rate along a trajectory

Consider the scenario in Figure 5 of an orphaned point source located along a straight road. If the detector vehicle is moving along the road, then the observed count rate will change proportional to $e^{-\mu r}/r^2$, as seen in Equation 3.13. It is typical to let the detector acquire counts during a time interval $[t, t + t_{acq}]$ before recording the data to the digital system.

Let $|r(t)|$ describe the Euclidean norm between the detector at time t on the road, and $\theta(t)$ the angle between $r(t)$ and the direction of travel. For anisotropic detectors, the count rate will depend on that incidence angle $\theta(t)$. Assuming a fixed velocity v, the total count rate acquired during the acquisition window is then

$$N(t) = \int_t^{t+t_{acq}} \underbrace{\dot{N}(r(\tau), E_\gamma, \theta(\tau))}_{\text{CPS}}\, d\tau + \text{Pois}(\lambda_{bkg}) \tag{3.14}$$

Numerically, the trajectory is discretized into a set of points uniformly spaced along the path length. These are mapped to time via $t = s/v$, after which the count rate is integrated over each acquisition interval using the trapezoidal rule.

3.3 To Bayes or not to Bayes

An analytical approach to source localisation, like the one described in Section 2.5, can be useful as a first estimate. In practice, the assumptions under which the analytical approach is taken are often violated. Bayesian statistics offers a framework to maximally utilize available data. Bukartas et al. [15] and Dvornik et al. [16] have shown that it is technically feasible to use Bayesian statistics for mobile search scenarios. Consider the foreword of *Bayes Rules! An Introduction to Applied Bayesian Modeling*, where Johnson et al. [30] describe the following advantages of Bayesian methods:

> "Though frequentist and Bayesian methods share a common goal – learning from data – the Bayesian approach to this goal is gaining popularity for many reasons: (1) Bayesian methods allow us to interpret new data in light of prior information, formally weaving both into a set of updated information; (2) relative to the confidence intervals and p-values utilized in frequentist analyses, Bayesian results are easier to interpret; (3) Bayesian methods can shine in settings where frequentist "likelihood" methods break down; and (4) the computational tools required for applying Bayesian techniques are increasingly accessible."

Each of these four points relates to mobile search scenarios.

1. Bayesian inference asks what is plausible given limited data and a known structure, whereas frequentist inference asks what is supported by the data alone. That is, instead of merely finding optimal parameters based on maximum likelihood,

$$\underbrace{p(y \mid \theta)}_{\text{sampling distribution}} \tag{3.15}$$

   Bayesian inference seeks to find

$$\underbrace{p(\theta \mid y)}_{\text{posterior inference}} \propto \underbrace{p(y \mid \theta)}_{\text{sampling distribution}} \times \underbrace{p(\theta)}_{\text{structural information}} \tag{3.16}$$

   which is a distribution of model parameters θ given the observed data y. The structural information takes the form of a *prior distribution* over the parameter space. This distribution does **not** depend on the data, but rather defines the prior *belief* about the possible parameters. It is possible to define *generic* priors, meaning there is no strong belief about the likely model parameters. However, physical laws, environmental boundaries or expert knowledge can be used to *inform* the choice of priors. In the context of source localization, this could be in the form of a restricted search area based on available information, or an estimate of the activity of the source. These could be based on external knowledge or a first measurement [16]. Informed priors are a rigorous probabilistic way to include prior information in the inference process.

2. Frequentist statistics computes the likelihood $p(y \mid \theta)$, which answers: *given this set of parameters θ, what is the probability that we observe this data y?* Bayesian inference provides a posterior $p(\theta \mid y)$, that answers: *given this set of data y, what is the probability that this data results from the parameters θ?* To link the former back to conclusions about the parameters requires some mental effort, whereas the latter Bayesian approach gives more direct insights, including credible intervals for estimated parameters. Linking to the case of source localization, the posterior $p(\theta \mid y)$ directly provides a probability distribution over the possible source locations and activities. Having the ability to directly quantify uncertainties is very beneficial for source localization, as e.g. false negatives could potentially have severe consequences.

3. Mobile gamma spectrometry typically yields sparse and noise data due to the limited counting statistics in an acquisition window and the distance between the source and the detector, often yielding a very weak signal in the count rate function. Bayesian methods can mitigate this, assuming appropriate priors are chosen [31].

4. The last few years have seen great advancements in computational power and software tools available for Bayesian statistics. Libraries like PyMC [32] have developed into comprehensive Bayesian libraries, suitable for defining a Bayesian search model and sampling the posteriors using Markov Chain Monte Carlo. for use with Python. Dvornik et al. [16] have shown a practical implementation of a Bayesian mobile search algorithm can be defined and sampled in PyMC well.

4 The PG-RAD software

4.1 Outline

This thesis introduces *PG-RAD - Primary Gamma RADiation Simulator, v0.1.0* [33], a command-line software package for simulating localisation of orphan sources using mobile gamma spectrometry. PG-RAD provides a framework for construction road geometries, arbitrary distributions of point sources, modelling the detector response to primary gammas. PG-RAD provides a configurable framework for constructing detector trajectories, defining radioactive source distributions, modelling detector response, and generating synthetic acquisition data. User input is specified through YAML configuration files, which makes simulations reproducible and easily shared. A visual overview of the software architecture is shown in Figure 7. The four main stages are:

1. The YAML configuration is parsed and validated into structured specification objects.
2. A simulation landscape is constructed, including the path geometry, source placements and detector specifications.
3. The simulation engine numerically approximates detector response as it travels along the trajectory.
4. Results are then post-processed, visualized and optionally saved to disk.

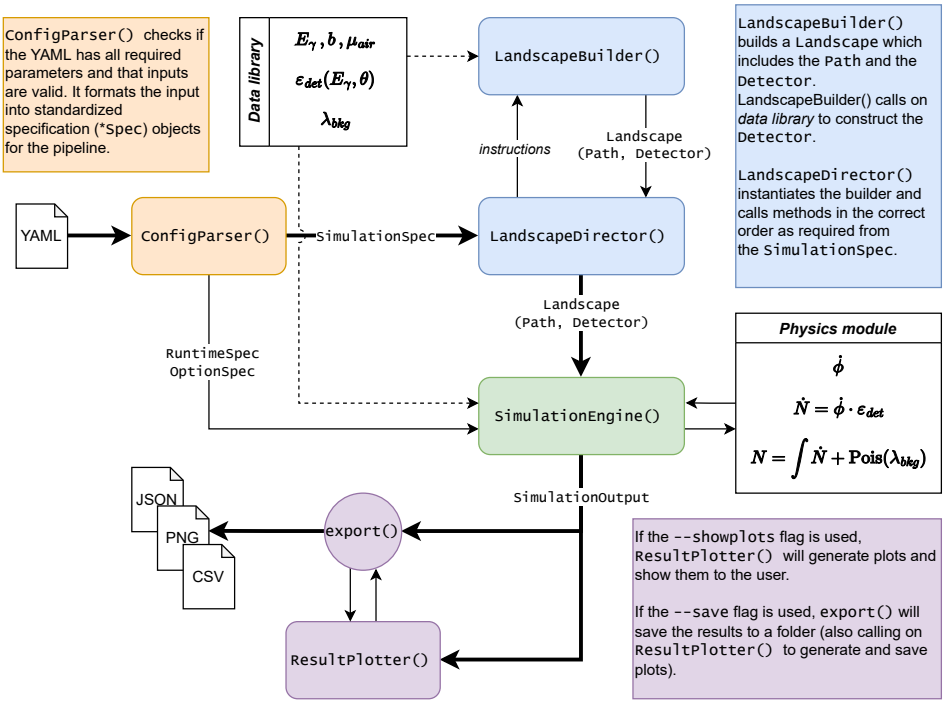


Figure 7: Schematic overview of PG-RAD modules and their interactions. The bold lines show the core sequence of operations that occurs when the user runs `pgrad -config input.yml`.

4.2 Raison d'être

Other data synthesizers have been used by the Medical Radiation Physics group of Lund University to simulate mobile gamma spectrometry scenarios. However, PG-RAD addresses the need for a flexible, user-friendly tool to generate diverse mobile gamma spectrometry scenarios. Compared to previous data synthesisers in use, PG-RAD offers the following:

- PG-RAD allows for placement of sources and trajectories in three dimensions.
- Place several point sources anywhere in the 3D space.
- Geometry of the trajectory can be defined using segments (e.g. straight segment, left turn segment), or be loaded in from experimental GPS data.
- PG-RAD is designed using object-oriented programming principles, ensuring that the code is extensible in the future. This will be discussed in more detail in Section 4.4.

4.3 Input and output

The structured input cards are written in YAML. YAML works on the principle of key-value pairs, and allows a high degree of flexibility in terms of data types and embedded key-value pairs. This declarative approach means that inputs are written in an exact format, producing the same output for the same input. Listing 1 shows an example of a complete, correct input to PG-RAD 0.1.1. The corresponding output is shown in Figure 8.

```yaml
name: Example landscape
speed: 8.33
acquisition_time: 1
path:
  length: 1000
  segments:
    - straight
    - turn_left: 90
    - straight
sources:
  s1:
    activity_MBq: 1000
    isotope: Cs137
    gamma_energy_keV: 662
    position:
      along_path: 300
      dist_from_path: 72
      side: left
  s2:
    activity_MBq: 1000
    isotope: Cs137
    gamma_energy_keV: 662
    position: [100, 100, 0]
detector: LU_HPGe_90
options:
  bkg_cps: 15
```

Listing 1: An example input file to PG-RAD, version 0.1.1. The output from this input file is shown in Figure 8. Notice in this case position definitions are mixed (relative and absolute positions), and the length is randomly distributed among the three segments. For more details on each parameter and possible options, see Appendix B.

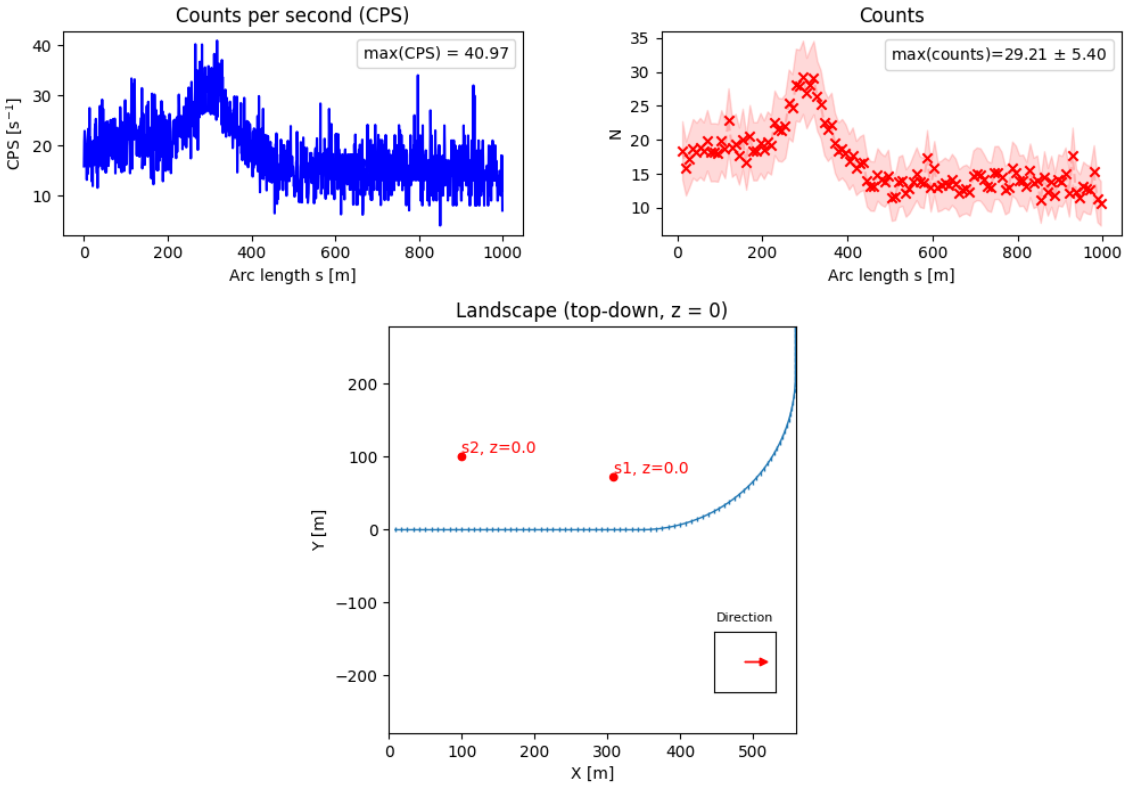


Figure 8: Plots produced by PG-RAD for the input shown in Listing 1. Note that the integrated counts after 1 second acquisition time are lower than the maximum of the count rate function. This has to do with alignment of integration windows as shown in Figure 6.

4.4 Design philosophy

The overall approach to designing PG-RAD was guided by four principles: 1) flexibility; 2) maintainability; 3) reproducibility; and 4) extensibility.

Flexibility

- Users can define arbitrary source constellations (point sources, activities, isotopes).
- Complex road geometries can be constructed from pre-built segments (e.g., straight, turn) or recreated from experimental GPS data.
- Sources can be aligned or misaligned with acquisition points (see Figure 6).
- Background radiation can be generated internally or manually specified by the user.
- Angular detector efficiency dependency can be toggled on/off.

Maintainability PG-RAD is written in Python 3 and takes advantage of the broad offering of high quality scientific packages available. In particular, PG-RAD uses NumPy [34, version 2.4.1] and SciPy [35, version 1.15.0] for interpolation and integration routines. Relying on well-tested libraries avoids unnecessary complexity in the code and reduces risk of implementation errors.

(a) Simulation parameters

Parameter	Value
Air density (kg/m^3)	1.243
Total path length (m)	999.6
Readout points	120
Seed	440869
Mean background (cps)	14.964

(b) Point sources

Name	Isotope	Activity (MBq)	Position (m)	r_{min} (m)
s1	Cs137 (661.7 keV)	1000.0	(308.33, 72.00, 0.00)	72.0
s2	Cs137 (661.7 keV)	1000.0	(100.00, 100.00, 0.00)	100.0

Table 1: Tabulated PG-RAD parameters based on the scenario in Listing 1 and Figure 8.

The code is structured in an object-oriented and modular way, separating geometry, physics, and data handling functionalities. Unit testing has been applied to core components, such as validation of the photon fluence calculation. These software engineering practices help to avoid building up technical debt and improves long-term maintainability of PG-RAD [36].

Extensibility Related to the previous point, PG-RAD has a separate data library which contains isotope information and detector response curves. These being separated from the core simulations ensures that new isotopes and detectors can be added by simply extending the data library, requiring no modification of the core PG-RAD code. The architecture could allow for extending

Reproducibility Reproducibility is an important aspect of scientific software. For this reason, user input takes the form of YAML input files. The advantage is that explicit input files allow for sharing and adapting in a reproducible manner. Using the same input file with the same version of PG-RAD will then produce the same output. Although the physics modelling is deterministic, the background activity (and possibly trajectory geometry, depending on user input) use samplers. To ensure identical output between runs, a global seed can be set in the input file to fix the samplers' behaviour.

4.5 Architecture

PointSource geometry The PG-RAD landscape can be populated with an arbitrary number of isotropic point sources which are constructed using the `PointSource` object. Under the `sources` keyword in the configuration, the user specifies a point sources with a unique subkey, which acts as their name and identifier. A fully configured source looks like:

```yaml
sources:
  example_source_1:
    activity_MBq: 1000
    isotope: Cs137
    gamma_energy_keV: 662
    position:
      along_path: 300
      dist_from_path: 32
      side: left
```

`activity_MBq` is the activity of the point source. The `isotope` is assigned by a string (e.g. Cs137, Co60, Cs134), and the user also specifies the primary gamma energy that is to be simulated under `gamma_energy_keV`. Note that this gamma energy does not need to be exact. PG-RAD has an internal data library which contains various known isotopes, their primary gamma lines E_γ, branching ratio b and μ_{air} The combination of the `isotope` and `gamma_energy_keV` keywords will retrieve all of this information. As of version 0.1.1, the included isotopes are ^{137}Cs, ^{134}Cs and ^{60}Co, but the library can be easily expanded in future updates to include other industrial isotopes such as ^{192}Ir, ^{241}Am or ^{91}Sr.

The position of the source can be specified relative to the path as in the above code block. Alternatively, absolute (x, y, z) coordinates can be provided. PG-RAD supports placing the sources anywhere in a 3D Cartesian space. There is no arbitrary limit to the size of the landscape. However, sources which are placed very close or on the road are handled by a minimum distance r, as to avoid numerical overflow issues due to division by r^2.

```
  position: [256, 321, 0]
```

This flexibility allows the user to put one or more sources in any arbitrary points in space, letting the user easily generate data for single or multi-source search scenarios.

Path geometry A path is specified in one of two ways. The first option is to specify a path using segment building blocks.

```
  path:
    length: 1000
    segments:
      - straight
      - turn_left: 90
      - straight
```

The currently supported segments are `straight`, `turn_left` and `turn_right`. One can specify the length as a single number, to be randomly distributed amongst the segments, or as a list of the same length as the segment list. Turn segments can optionally get a specified angle in degrees. If left out, this too will be randomly sampled. For an explanation on how the random sampling works, see Appendix A.

Detectors A data library is set up in the PG-RAD package, which allows `LandscapeBuilder()` to retrieve detector information and add it to the `Landscape`. There are currently three detectors set up. The first is an isotropic (not angularly dependent) 3×3 inch NaI(Tl) detector (where 3×3 inch refers to the sensitive detecting volume of the detector). Second, two angular detectors are setup: a 123% efficiency HPGe detector and a 4 L NaI(Tl) detector. The efficiency response for these detectors is based on calibrations performed by the team of Medical Radiation Physics at Lund University, and are thus specific to their setup. The data library stores the name of the detector, the type of the detector (NaI(Tl) or HPGe), the field efficiency as a function of incoming photon energy and the relative angular efficiency relative to the direction of travel (both in tabulated form). Similarly to the sources, it would be straightforward to add other detectors in the future; all that is needed is the (angular) detector efficiency as a function of photon energy.

As discussed in Section 2.4.4, the relative angular efficiency of anisotropic detectors is particular to the setup in which the detector is placed. These tables can be accessed by the `Detector` object when needed to calculate the count rate per second in the `SimulationEngine()`. Because the efficiencies are tabulated values which were experimentally determined using

calibration sources, the efficiency is interpolated if the incoming photon energy falls between any two tabulated values.

As of PG-RAD version 0.1.1, it is not possible to specify a custom detector. However, because data is retrieved from a separate library in order to build the `Detector` object, it is straightforward to add more detectors to the data library in the future, allowing PG-RAD to use other detectors.

Landscape building The landscape is built using a set of instructions which are given by the `LandscapeDirector()` to the `LandscapeBuilder()`. The `LandscapeDirector()`, as its name implies, directs the method and order that the `LandscapeBuilder()` should construct the landscape. The reason for this design is not arbitrary. Rather, it follows the so-called *builder design pattern* [37, 38], which is intended to separate the construction of a complex object from its particular configuration. This pattern is suitable for landscape construction, as it involves various independent components.

Simulation Engine The `SimulationEngine()` is responsible for producing the numerical results. Given a constructed landscape, it will perform these steps:

1. Create artificial subpoints between each acquisition point along the trajectory, and simulate the photon fluence at each of them following Equation 3.11.
2. The fluence is scaled by the efficiency. If the user specified an angularly dependent detector, this will be accounted for.
3. Background radiation is added to the count rate, which is either sampled from the internal library or based on the user-provided mean background count rate.
4. The trajectory is mapped to the time domain and integrated as in Equation 3.14 using the trapezoidal rule, producing discrete acquisitions.
5. Some diagnostic quantities are computed and processed for later analysis.

5 Application of PG-RAD

5.1 Validation

The first application of PG-RAD is to do validation of the outputs. This is done in two ways. First, a comparison is drawn with two existing in-house codes, developed at the Institute of Translational Medicine, Lund University. To see whether the count rate function is similar in PG-RAD compared to existing codes, a single-source configuration is constructed in all three codes. A parameter sweep is then done across a range of source activities and distances to the detector.

The second validation test is to recreate experimental scenarios in PG-RAD. PG-RAD allows the user to use a CSV file containing experimental coordinates as a path, thus enabling easy recreation. Data from an October 2024 mobile search field experiment [16] and performance is assessed by comparing the counts along the trajectory. The goal of this test is to see whether PG-RAD is able to generate data that represents real-world scenarios.

5.2 3D scenario

PG-RAD is capable of simulating basic 3D constellations of source and path geometries, which is an improvement over current data synthesizers. In order to demonstrate this capability, a mobile search scenario representing a drone carrying a 3×3 inch NaI(Tl) detector is simulated. This demonstration is relevant because the field of UAV-based gamma spectrometry is an active field of research. The aerial view of the gamma field can be complimentary to other measurement techniques when it comes to locating lost sources or informing clean up of nuclear fallout [39, 40]. Applying Bayesian methods to such spatial data is also an area of investigation [27], but this will not be tested in this thesis.

5.3 Testing the car-based Bayesian localisation method

5.3.1 Source geometries

A two source Bayesian localisation algorithm is defined based on the methodology of Dvornik et al. [16] and is shown in Figure 9. In order to test this model, the source geometry is varied along two parameters. The first test will be to vary the separation distance d of two identical sources. The second test will be to change the ratio of activities A_1/A_2 of the sources.

5.3.2 Performance metrics

To evaluate the localisation performance of the Bayesian model, a few quantities need to be specified.

Activity Relative Deviation The ARD is defined as the relative difference between the estimated activity of a source and its true activity in percent:

$$ARD = \frac{|A_{est} - A_{true}|}{A_{true}} \cdot 100 \tag{5.1}$$

Relative Position Deviation The RPD is defined as the relative difference between the true minimum detector-to-source distance and the estimated one. Similarly to the ARD:

$$RPD = \frac{|r_{est} - r_{min}|}{r_{min}} \cdot 100 \tag{5.2}$$

Localisation probability Dvornik et al. [16] define a successful source localisation as a posterior estimate of activity and location of the source such that both the ARD and RPD are below a threshold of 30%. This threshold was chosen, because posterior estimates with a higher ARD or RPD than this threshold were observed to have significant uncertainties and their interpretation is therefore ambiguous [16].

In this thesis, 10 model runs were performed for each test scenario, resulting in $2 \times 10 = 20$ localisation attempts for which the ARD and RPD are evaluated. The *localisation probability* (LP) is then the percentage of successful localizations and the total number of runs performed by the model.

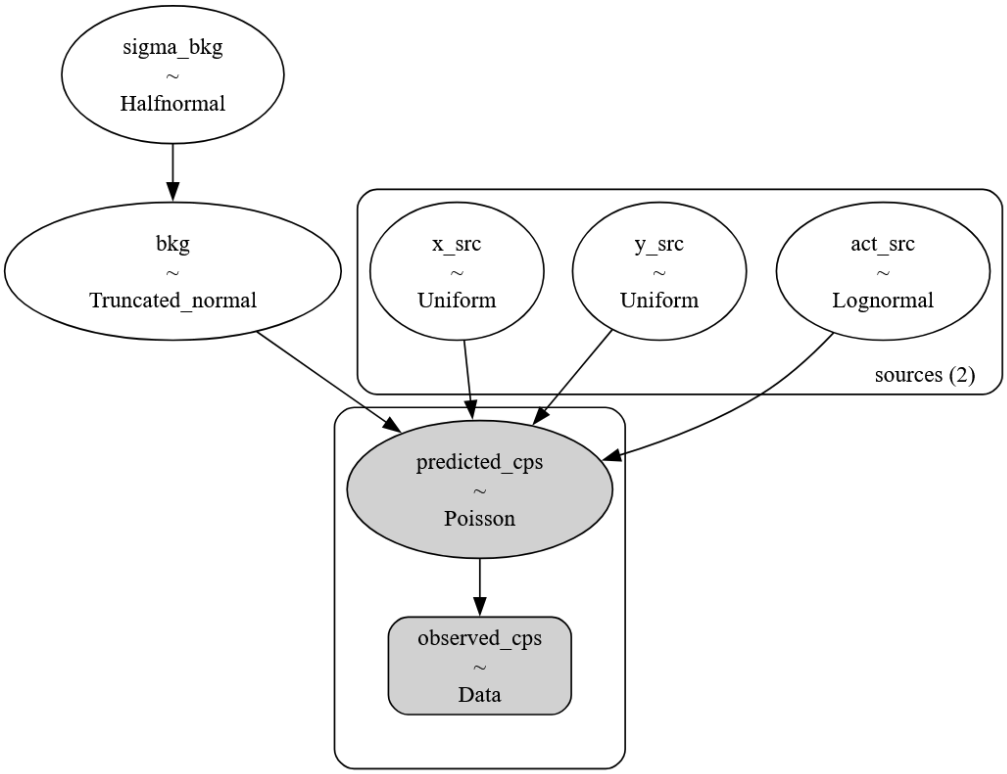


Figure 9: Schematic representation of the two-source Bayesian model. The search area for the location is defined as $S = [a, b] \times [c, d]$. An uninformed prior would then be $x_{src} \sim \text{Uniform}(a, b)$ and $y_{src} \sim \text{Uniform}(c, d)$. Assuming the road is not perpendicular to either axis, the drawing a square area $S = [x_{min}, x_{max}] \times [y_{min}, y_{max}]$, where (x_{min}, y_{min}) and (x_{max}, y_{max}) are the beginning and end coordinates of the road respectively, is a reasonable search area. This was the approach used in the consequent analysis. The activity is modelled as $act_{src} \sim \text{LogNormal}(\mu, \sigma)$, as this distribution is suitable for values which span several orders of magnitude. The background activity is $bkg \sim \text{TruncatedNormal}(\mu, \sigma)$ where $\sigma \sim \text{HalfNormal}(\sigma_{data})$. Here, σ_{data} is the variance of the background counts, and μ the mean of the background counts. The TruncatedNormal distribution is used for the background activity so that the left part of the distribution is truncated at 0 (no negative counts). Finally, the predicted counts are modelled as a Poisson distributed variable, with the physical model identical to that in Equation 3.14. In the tests performed in this thesis, the background counts are kept constant when generating PG-RAD data for the Bayesian model, and thus they are not sampled but included as a constant level added to the net counts.

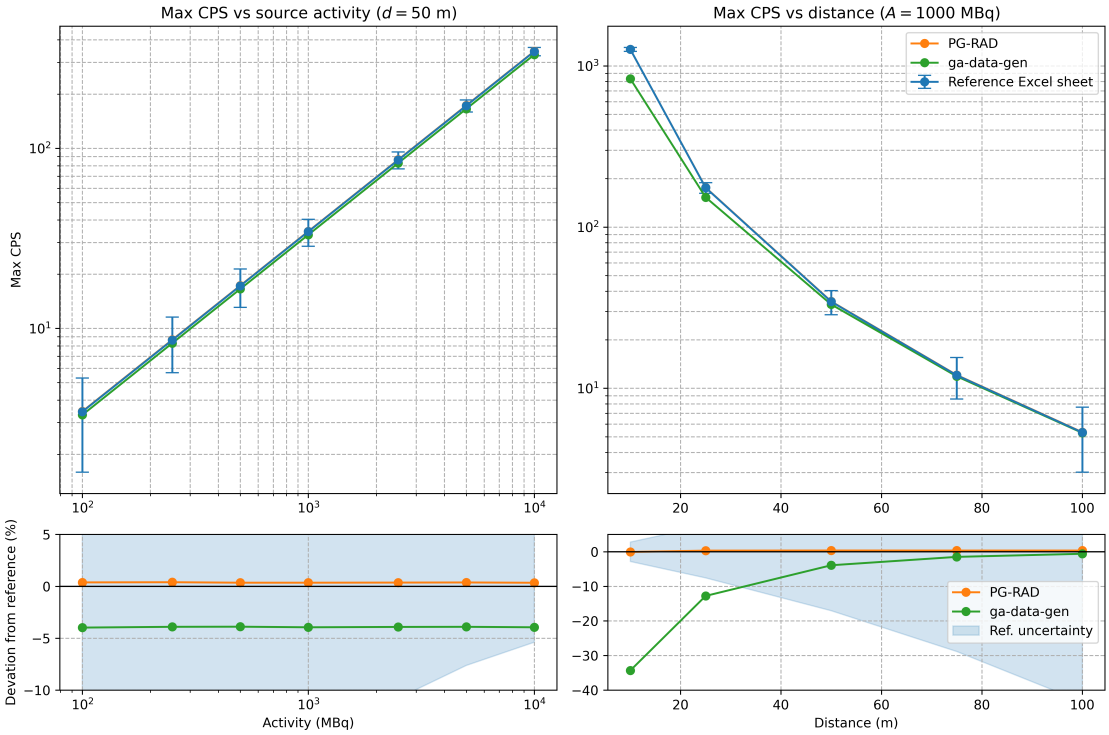


Figure 10: PG-RAD compared to two other codes used for gamma spectrometry calculations. All three codes use the same basic physical model as described by Equation 3.11 and Equation 3.13, and also use the same calibration data for the detector efficiency. Max CPS refers to the maximum count rate per second observed in the count rate function. How the max CPS changes with respect to source activity and distance is shown in the top row. The bottom row shows the relative deviation between the codes, using the Excel routine as a reference. The uncertainty band represents the $\sigma = \sqrt{N}$ which comes from the counting statistics. Overall, PG-RAD is in good agreement with the reference Excel routine.

6 Results and discussion

6.1 Validating PG-RAD output

6.1.1 Comparison to other in-house codes

To verify the physical accuracy of PG-RAD, an identical scenario was created in PG-RAD and two other in-house codes: A single Cs-137 source measured by a 3×3 inch NaI(Tl) detector, with no natural background activity. Two variables were then varied: the distance to the road r_{min} and the activity of the source A. To compare the output of the three codes, the CPS recorded at the peak was used as a benchmark.

The purpose of this test whether the entire pipeline works: whether input is correctly interpreted by `InputParser()`, `LandscapeBuilder()` correctly builds the expected landscape, and `SimulationEngine()` correctly interpolates, integrates and generates the outputs. In short, all the parts of PG-RAD will be involved. This is a first diagnostic to see whether there are any structural errors in PG-RAD as compared to a reference code which has been validated. Figure 10 shows that there is good agreement between PG-RAD and the reference Excel routine, when varying both source distance and activity.

6.1.2 Recreating experimental scenarios

ID	Length, km	Source act., MBq	bkg. CPS (4L NaI(Tl))	r_{min}, m	Side
Setup 1	0.36	1251 ± 41	51 ± 9	82	Right
Road 2	77.6	565 ± 18	91 ± 15	100	Right
Road 3	71.9	601 ± 20	90 ± 17	72	Left

Table 2: Parameters of October 2024 field search experiments using a ^{137}Cs point source.

```yaml
name: Setup1_F15
acquisition_time: 1
detector: {DET}
speed: 4.16
path:
  file: [...]
  east_col_name: East
  north_col_name: North
  z: 1
sources:
  s1:
    activity_MBq: 1251
    gamma_energy_keV: 662
    isotope: Cs137
    position:
      along_path: 170
      dist_from_path: 77
      side: right
options:
  bkg_cps: 0
```

(a) "Setup 1"

```yaml
name: Road2_F30
acquisition_time: 1
detector: LU_NaIR
speed: 16.67
path:
  file: [...]
  east_col_name: East
  north_col_name: North
  z: 1
sources:
  s1:
    activity_MBq: 565
    gamma_energy_keV: 662
    isotope: Cs137
    position:
      along_path: 550
      dist_from_path: 100
      side: right
options:
  bkg_cps: 106
```

(b) "Road 2"

```yaml
name: Road3_F30
acquisition_time: 1
detector: LU_NaIR
speed: 16.67
path:
  file: [...]
  east_col_name: East
  north_col_name: North
  z: 1
sources:
  s1:
    activity_MBq: 601
    gamma_energy_keV: 662
    isotope: Cs137
    position:
      along_path: 495
      dist_from_path: 72
      side: left
options:
  bkg_cps: 107
```

(c) "Road 3"

Figure 11: YAML input files reconstructing three experimental scenarios, following Table 2 as described by Dvornik et al. [16]. Note that for gross counts in 'Road 2' and 'Road 3', the upper estimate of the background counts was used. In "Setup 1", `{DET}` is placeholder for `LU_NaIR` or `LU_HPGe_90`, as both of these were simulated. Filenames are omitted for brevity.

Three measurements used the 4 L NaI(Tl) detector that is implemented in PG-RAD, and one using the HPGe detector currently implemented in PG-RAD. The ground truth information on source activity and location is displayed in Table 2. This information was used to create PG-RAD input files, which are shown in Figure 11. The resulting generated counts along the path are compared to the experimentally observed counts in Figure 13.

PG-RAD manages to recreate the HPGe data of 'Setup 1' rather well. The background corrected experimental counts and the simulated counts without background show a similarly shaped peak and a peak number of counts within the margin of error. On the left side of the peak, experimental counts are slightly lower than PG-RAD, which is due to light vegetation shielding the source there. On the right side, the peak shows an identical sharp drop. Looking at the visualisation of the scenario in Figure 12, one can see the asymmetry comes from the change in direction of the trajectory right after passing the source, causing a distorted asymmetric peak compared to a straight road trajectory. Regarding the NaI(Tl) data, PG-RAD recreates the same curve, but with a significantly lower number of counts. For the scenarios

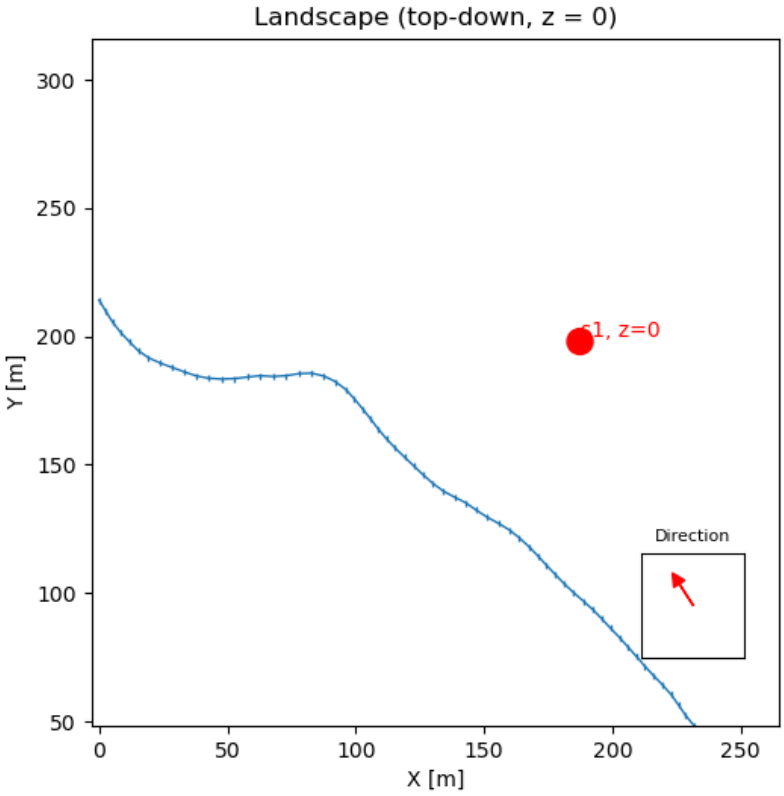


Figure 12: The experimental scenario 'Setup 1'. The resulting counts along the path can be seen in Figure 13a.

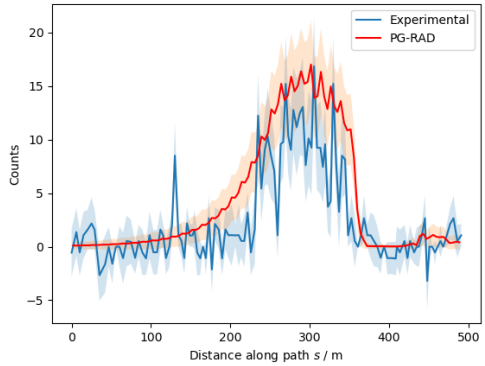
'Road 2' and 'Road 3', PG-RAD once again shows good agreement with experiment when it comes to the shape of the curves, but appears to underestimate the counts in the 'Road 2' scenario.

NaI(Tl) detectors, having a poorer energy resolution than HPGe detectors, will count some Compton scattered photons in the same energy bin as the selected ROI (see Section 2.4). PG-RAD does not simulate Compton scattering effects, and this could in part explain the higher experimental count rates observed in 'Setup 1' and 'Road 2'. In 'Road 3' however, which also uses the same NaI(Tl) detector, there is no such discrepancy. This could be explained by the experimental conditions. 'Setup 1' and 'Road 2' involved little to no shielding, with the first having minor vegetation, and the latter source being in an open field. In contrast, the source in scenario 'Road 3' was located in a forest next to the road, meaning considerable shielding was present (see Dvornik et al. [16] for further experimental context). When shielding objects are present in the gamma field, there is a decrease in primary gamma photons and forward scattered Compton photons that reach the detector. PG-RAD and experimental data thus likely line up better only because the shielding from the forest is reducing the count rate in the experimental data. If there were no forest shielding the source, a discrepancy similar to that of the other two scenarios could be expected between PG-RAD and experimental data.

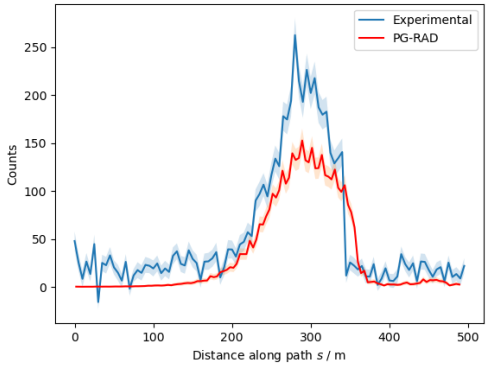
HPGe detectors will on the other hand essentially only record primary gammas in the region of interest. This further supports the idea that Compton scattering plays a primary role in the discrepancies observed between PG-RAD and experimental data. could explain why, in the HPGe data of 'Setup 1', the PG-RAD simulated HPGe counts are slightly higher than experi-

mental (although within margin of error) and the NaI(Tl) counts are lower. However, other sources of error could be present besides merely the lack of Compton photons in the ROI. Further investigation should be done to identify such possible errors (e.g. by recreating more scenarios involving NaI(Tl) detectors).

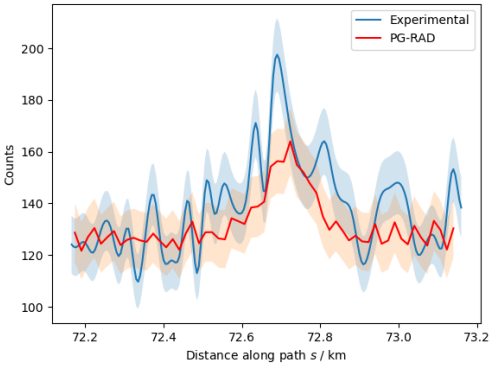
Finally, regarding the gross count rate simulations, one can see that PG-RAD simulated background, which is simulated as a Poisson process with constant average CPS, has a lot less variation than the true background counts. This has two possible explanations. First, the Poisson background is applied to the CPS prior to integration. This will inevitably lead to some smoothing in the background. Second, the assumption of a constant mean for the background count rate is not valid in real-life, as the the natural background can fluctuate along the trajectory.



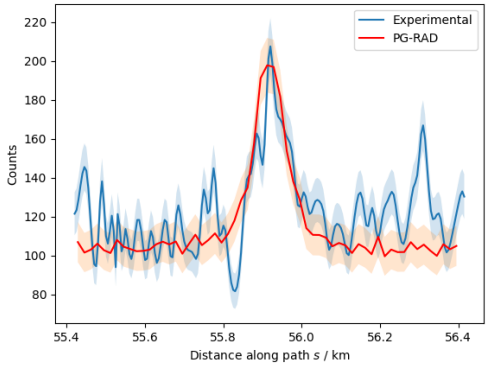
(a) "Setup 1", HPGe (net counts)



(b) "Setup 1", NaI(Tl) (net counts)



(c) "Road 2", NaI(Tl) (gross counts)



(d) "Road 3", NaI(Tl) (gross counts)

Figure 13: PG-RAD compared to experimental data of the October 2024 field experiment as described in Dvornik et al. [16]. The uncertainty bands are an estimate representing $\sqrt{\text{counts}}$.

6.1.3 3D scenario - UAV search

Although complex three-dimensional paths are not possible in PG-RAD as of version 0.1.1, there is the possibility to set the path or sources to a certain height level z, thus making PG-RAD capable of simulating UAV search scenarios. Figure 14 shows a possible scenario where a UAV equipped with a 3×3 inch NaI(Tl) detector searches a rectangular area in a scanning

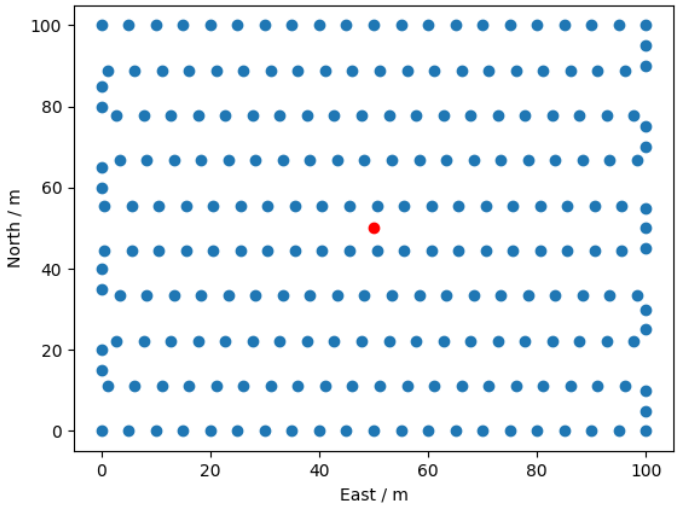
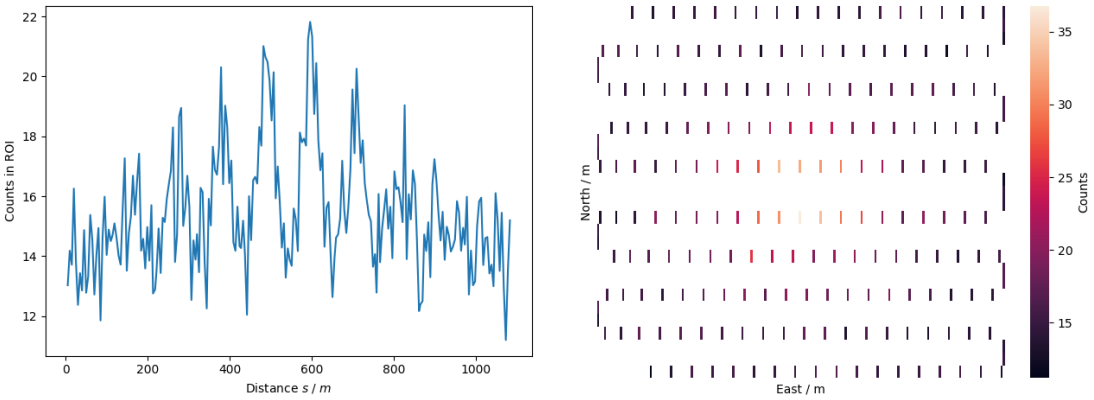


Figure 14: The UAV scenario and resulting PG-RAD counts along the path. A 3×3 inch NaI(Tl) detector is travelling at $5\,\mathrm{m\,s^{-1}}$ at a fixed altitude z, searching a $100\,\mathrm{m}^2$ area. The acquisition time t_{acq} is 1 s. The source is a 50 MBq ^{137}Cs source located at (50, 50).

pattern. The path itself was generated outside of PG-RAD, and then passed onto the program as a CSV file. PG-RAD can then calculate the expected counts along the path, as shown in Figure 15. The same figure also shows the counts projected back to the (x, y) plane.

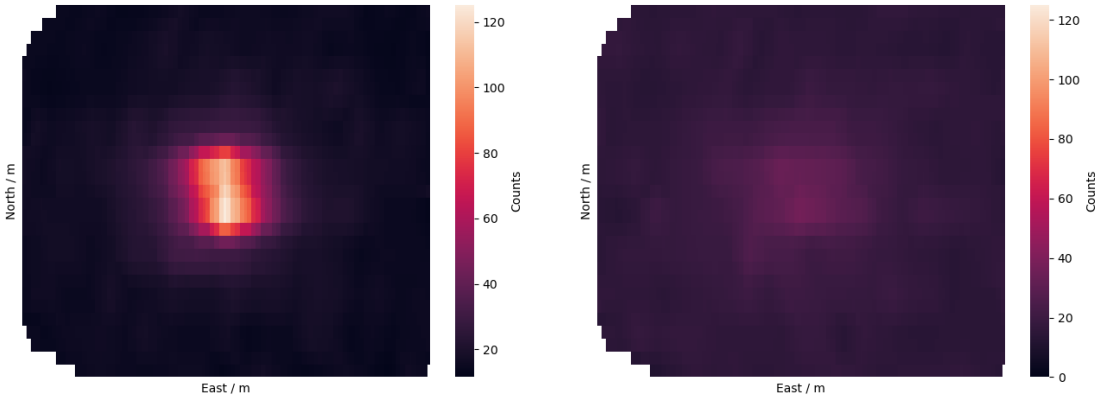
Figure 16 shows a linear interpolation of the acquired counts. The result is a heat map showing the spatial distribution of the gamma field, simulated from PG-RAD. In particular, one can notice the challenge of background radiation in a UAV search scenario as the altitude is increased. This scenario was created in order to show that PG-RAD functions in three dimensions, and is thus capable of being used to sketch basic UAV-based source localisation scenarios.



(a) Counts along trajectory

(b) Spatial mapping

Figure 15: The counts along the trajectory at $z = 15\,\mathrm{m}$. This is the output directly generated by PG-RAD. b) Counts projected back to the (x, y) plane.



(a) $z = 5\,\mathrm{m}$ (b) $z = 25\,\mathrm{m}$

Figure 16: Post-processed and interpolated spatial map of the counts, for a) $z = 5\,\mathrm{m}$ and b) $z = 15\,\mathrm{m}$ altitude. At higher altitudes, due to the inverse square relationship between fluence and distance to the source (Equation 3.11), there is a wider spread of photon counts, and they are less distiguishable from background radiation. The inverse square relation means resolvability of the location of the source significantly as a UAV is flown at higher altitudes.

6.2 Using generated data for testing Bayesian algorithm

Source separation Two identical 300 MBq ^{137}Cs sources at 32 m from the road are separated by varying multiples of the full width at half maximum (FWHM) if a single peak. Figure 17 shows a selection of separation distances. A Python script was written to dynamically generate PG-RAD input cards between 4 and 1 FWHM separation. In the same pipeline, the generated data is passed onto the Bayesian model. For each separation distance, 10 runs of the model were performed, in order to compute the localisation probability as defined in Section 5.3.2. The goal of this test is to find how close two sources can be before the model starts to fail localisation. Figure 18 shows the localization probability for this test case. The data was generated using PG-RAD and passed onto a pipeline performing 10 runs of the Bayesian model for each separation distance.

Localization Probability as a function of source activity ratio Two ^{137}Cs sources, with activities A_1 and A_2 are placed 32 m from the road are separated by 4 times the FWHM. As can be seen in Figure 19, the source strength A_2 is gradually reduced, such that the ratio A_1/A_2 increases. Similarly to the source separation test, the PG-RAD input was dynamically generated using a Python script and 10 runs of the model were performed. For this test, the goal is to see whether a significant difference in activity of two sources reduces localisation performance of the Bayesian model. Figure 20 shows the LP as a function of activity difference. Like the previous test, the data was generated dynamically in a Python script and passed onto the Bayesian pipeline. Once again, 10 runs were performed per configuration.

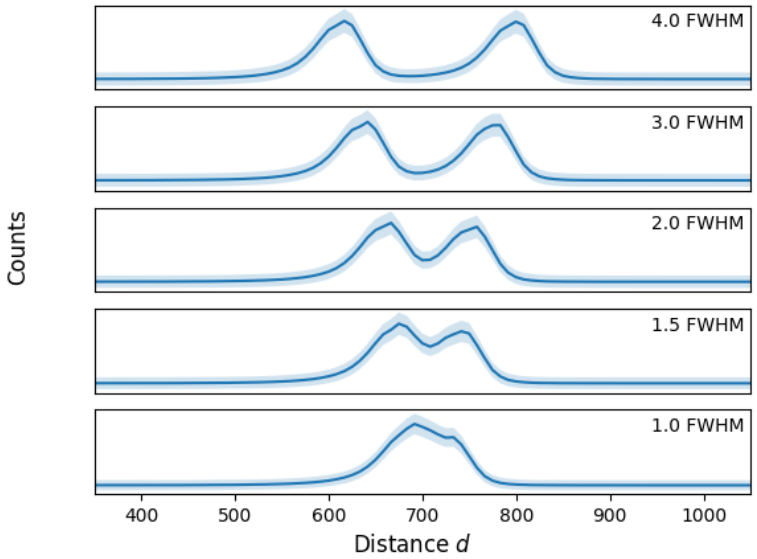
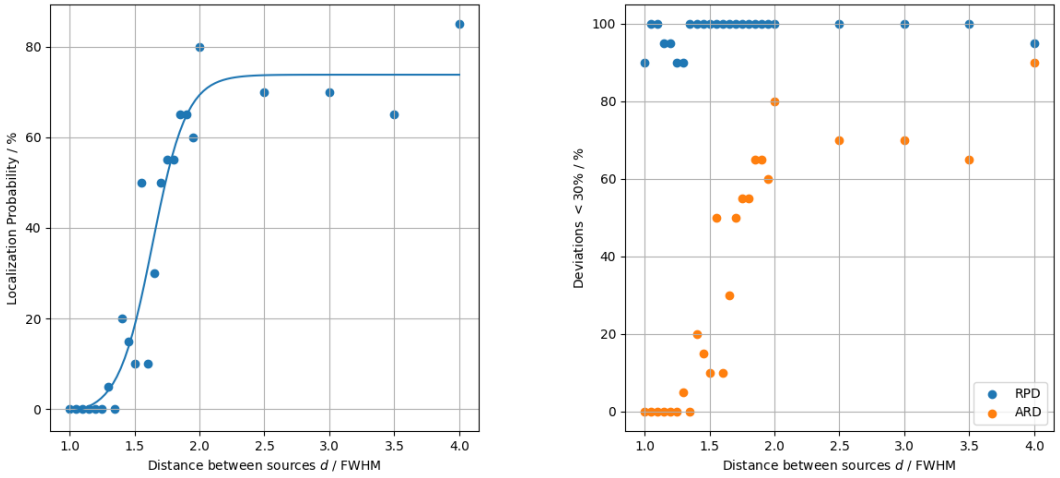


Figure 17: PG-RAD simulated counts for two 300 MBq ^{137}Cs sources at 32 m distance from a straight road, plotted for select source separations d. The spatial separation is expressed in multiples of the full width at half maximum of a single source peak.



(a) (b)

Figure 18: The localisation performance as a function of source separation d. The simulation used an HPGe detector travelling at 30 km h^{-1}, measuring two 300 MBq ^{137}Cs sources at 32 m from the road. Figure a) shows the overall localisation probability (LP) as defined in Section 5.3.2. A sigmoidal fit is performed to the LP, which shows that the LP goes below 50% at a source separation of around 1.6 to 1.8 FWHM. Figure b) shows the percentage of scenarios with an ARD and RPD below 30% as a function of source separation d. The ARD is the dominant factor in a reduced LP at small source separations d. The RPD remains below the threshold, showing in this particular setup, the two-source algorithm continues to be able to locate the two sources, while the activity estimate becomes inaccurate.

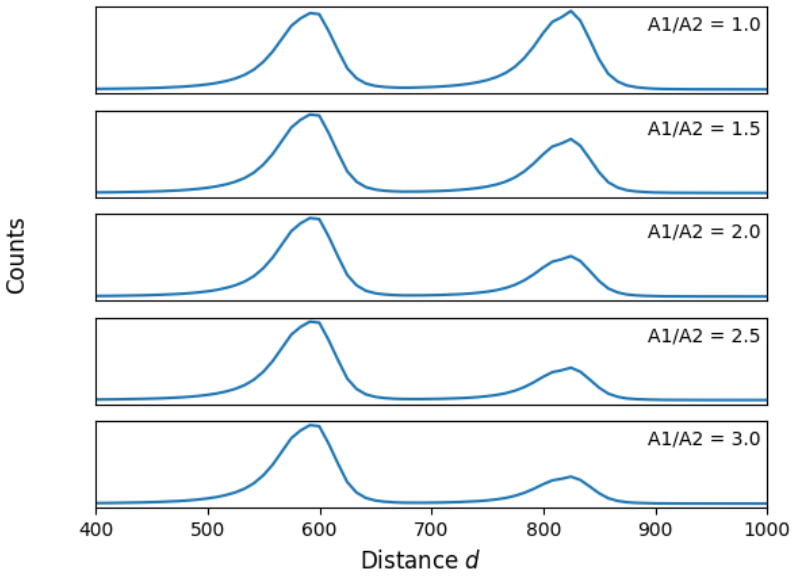


Figure 19: PG-RAD simulated counts for two ^{137}Cs sources at 4 FWHM separation plotted for select activity fractions A_1/A_2. A_1 is set to 300 MBq, and A_2 varying between 300 MBq and 100 MBq. When reducing A_2, we must make sure that the activity does not become so low that the source is undetectable. The distance from the straight road for both sources is thus set at 32 m, 15 m closer than the maximum detection distance of 47 m. This maximum detection distance is based on the results from Finck et al. [28] and are based on an unshielded ^{137}Cs source with an activity of 100 MBq, and is for the same HPGe detector which is also used in the PG-RAD simulations.

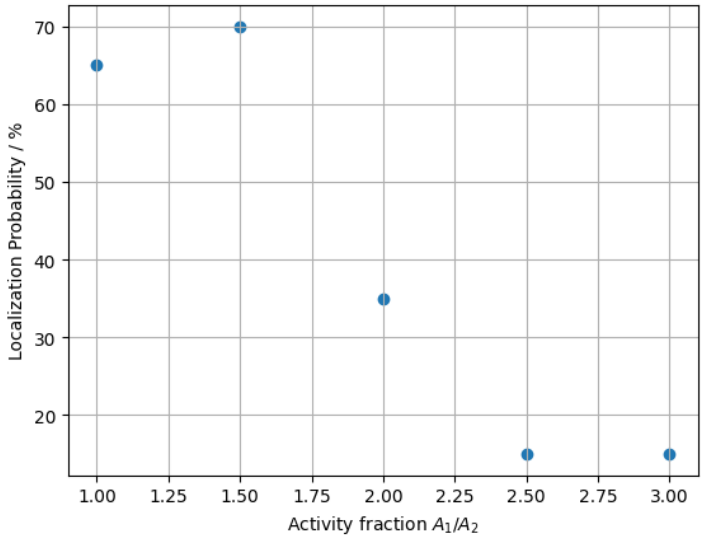


Figure 20: The localisation performance as a function of activity fraction. The simulation used an HPGe detector travelling at 30 km h^{-1}, measuring two ^{137}Cs sources at 32 m from the road. The localisation probability (LP) is as defined in Section 5.3.2. The LP appears to reduce rapidly to below 50% when A_1/A_2 approaches 2.

7 Conclusion

This thesis presented PG-RAD, a simulator designed to model mobile search scenarios using gamma spectrometry. The various scenarios generated show that PG-RAD is able to generate data with non-trivial source and path geometries in three dimensions. It is also able to mimic experimental data to a reasonable degree, although the lack of Compton scattered radiation in PG-RAD means the count rates are underestimated compared to real-life scenarios, which is particularly true for scintillator detectors. The lack of shielding presents another limiting factor in generating representative data. However, the focus on extendability and modular design for PG-RAD means adding such features is possible.

Regarding the Bayesian algorithm, PG-RAD has shown to be useful as a tool for validation of the algorithm. In particular, we showed how PG-RAD could be used to generate various scenarios. That is also possible to do dynamically, as was done for the testing of the Bayesian algorithm; PG-RAD input files were dynamically written and passed onto the Bayesian model. This shows PG-RAD is not just useful for generating single scenarios but can also be used to systematically generate test scenarios by updating parameters in the input file. Applying this generated data to the two-source Bayesian algorithm has given tangible insights to the limits of the two-source Bayesian algorithm as a function of the source separation and the activity difference between the two. The tests show that the localisation probability decreases as the two sources get closer, and likewise decreases as the ratio between activities of the sources increases.

In conclusion, PG-RAD is an early-stage development of a scientific toolbox for simulating mobile gamma spectrometry scenarios to support the development of new mobile search methods. The modular approach means PG-RAD could be improved in its physics or accommodate other types of source geometries. PG-RAD could also develop to a useful tool for other research teams around the world, by adding their characterised detector efficiencies and isotopes of interest to the library or allowing the user to specify detectors and isotopes as part of the input file.

8 Future outlook

Distributed area sources Some emergency scenarios involving radioactive MORC do not consist of just a single isolated point source. For example, in a nuclear fallout scenario, the MORC may be distributed over large areas. PG-RAD could be extended to include the possibility to define a 2D or even 3D distribution of radioactive material rather than merely point source geometries. The physics implementation does not need to be adjusted in a major way, as it already computes the distance vector in 3D Cartesian space. Rather, it is a matter of handling a user input. For example, a rectangular area source (flat in the (x, y) plane) could take the form of

```
sources:
  field_source_1:
    activity_MBq: 1000
    isotope: Cs137
    gamma_energy_keV: 662
    position:
      start: [0, 0]
      end: [50, 50]
```

Computationally, the field source can be approximated by points (grid-based approach), allowing PG-RAD physics library to handle the fluence and count rate calculations without any changes to those implementations. Alternatively, the physics library could be modified to include an analytical formulation for area sources (see e.g. Isaksson and Rääf [19, section 2.6.2]). In reality though, area sources are not necessarily uniformly distributed. PG-RAD would need to undergo significant changes to potentially allow the uploading of some CSV or array structure representing a non-uniform area source into the program.

Adding custom detectors/isotopes As of version 0.1.1, there is no user interface to add custom detectors or add new isotopes in PG-RAD. These can only be added by manually adding the appropriate CSV files into the data library, but this is not the intended workflow for end-users. Future work could include a user-friendly way to incorporate custom detectors, either by providing a command to add new detectors/isotopes.

Compton scattering and detector effects As the name PG-RAD suggests, the fundamental assumption in the physical model is only modelling primary gammas which deposit their full energy in the detector. However, as was seen in the results, secondary radiation such as Compton scattering is not negligible, particularly for scintillation detectors. Another effect that is not modelled is partial absorptions of primary gammas in the detector, which is primarily relevant for the lower energy range in scintillation detectors.

Other types of uncertainties In PG-RAD simulations, the standard deviation of $\sqrt{\text{counts}}$ which is due to the Poisson counting statistics is the dominant contributor to uncertainty in the estimated count function along the path. However, there are also other factors, which have not been included. For example, the characterisation of detector efficiency comes with uncertainty on the order of $\pm 3-5\%$, due to calibration source activity uncertainty and the experimental nature of the characterisation. Another smaller factor is the air density which is used to compute the air attenuation μ_{air}. In all scenarios, an air density average of $1.243\,\text{kg}\,\text{m}^{-3}$ was used, representative of Southern Skåne, Sweden. However, air density in Southern Skåne can vary up to $\pm 6\%$ throughout the year [4]. Propagating this uncertainty to the count rate will mean an additional uncertainty of a few percent. For experimental recreation, this source of error could be eliminated by using the exact air density of the day and location of the measurement from the records of local meteorological data.

Shielding There is currently no way to add shielding objects to PG-RAD. This is however a very common scenario. As could be seen in the recreation of experimental scenarios, the count rate is significantly affected by shielding. Future work should focus on first finding a way to add simple cubic structures made of a known material into the landscape. Then, the physics library and the `SimulationEngine` would need to be modified such that as the simulation along the path is performed, if there is an integration step where the source-to-detector vector passes through a shielding object, the attenuation in the fluence rate function is computed correctly. This is not trivial, but the framework is there for this feature to be added to PG-RAD.

More representative background radiation Currently, background radiation is sampled from a Poisson distribution which is assumed to have a constant mean value. As was seen in recreations of experimental scenarios, real-world variance in background is significantly higher. The integration is likely smoothing the background, and this could be improved by

applying background counts after the integration step, not before. However, there are further improvements to be made. For any mobile search along a section of road longer than a few hundred meters, the assumption of a constant mean of the background counts is not valid. This is because the background radiation varies with altitude as well as longitude and latitude. One option could be to treat the mean value λ of the background radiation as an autoregressive process, thus allowing for fluctuations that are more representative of real-life scenarios. Another aspect is that experimental data along highways has shown distinct peaks in count rate upon passing underneath viaducts. Such sharp discontinuities challenge the current implementation of the Bayesian localization algorithm. Allowing the user to add such features manually could further enhance the usefulness of PG-RAD in assessing localization algorithm performance.

References

[1] IAEA, *Nuclear Security Recommendations on Nuclear and Other Radioactive Material out of Regulatory Control: Recommendations* (IAEA Nuclear Security Series 15). Vienna: International Atomic Energy Agency, 2011, 48 pp., ISBN: 978-92-0-112210-0.

[2] IAEA, "IAEA Incident and Trafficking Database," IAEA, Vienna, Factsheet, 2026.

[3] Krane, *Introductory Nuclear Physics*. John Wiley & Sons, 1987, ISBN: 047180533X.

[4] R. R. Finck, "High Resolution Field Gamma Spectrometry and its Application to Problems In Environmental Radiology," 1992.

[5] R. Finck, K. Lidén, and B. Persson, "GAMMASPEKTROMETRISKA FÄLTSTUDIER II," Lund, LURI 1973-04, 1973.

[6] P. A. Aarnio, J. J. Ala-Heikkilä, T. T. Hakulinen, and M. T. Nikkinen, "Gamma spectrometric monitoring of environmental radioactivity using a mobile equipment," *J Radioanal Nucl Chem*, vol. 233, no. 1, pp. 217–223, Jul. 1, 1998, ISSN: 1588-2780. DOI: 10.1007/BF02389675.

[7] W. K. Gummer and Canada, Eds., *Cosmos 954: The Occurrence and Nature of Recovered Debris*. [Ottawa] : [Hull, Que: Atomic Energy Control Board ; available by mail from Canadian Govt. Pub. Centre Supply and Services Canada], 1980, 60 pp., ISBN: 978-0-660-10589-5.

[8] "Sources and effects of ionizing radiation. Annex J: Exposures and effects of the Chernobyl accident," United Nations, New York, 2000.

[9] K. Rissanen and T. Rahola, "CS-137 concentration in reindeer and its fodder plants," *Science of The Total Environment*, Transfer of Radionuclides to Livestock, vol. 85, pp. 199–206, Sep. 1, 1989, ISSN: 0048-9697. DOI: 10.1016/0048-9697(89)90318-5.

[10] IAEA. "Dosimetric and Medical Aspects of the Radiological Accident in Goiânia in 1987," Accessed: Mar. 11, 2026. [Online]. Available: http://www.iaea.org/publications/5305/dosimetric-and-medical-aspects-of-the-radiological-accident-in-goiania-in-1987.

[11] H. Belot, "Tiny radioactive capsule lost in Australian outback found by side of 1,400km stretch of road," *The Guardian*, Feb. 1, 2023, ISSN: 0261-3077.

[12] M. Holmberg, K. Edvarson, and R. Finck, "Radiation Doses in Sweden Resulting from the Chernobyl Fallout: A Review," *International Journal of Radiation Biology*, vol. 54, no. 2, pp. 151–166, Jan. 1988, ISSN: 0955-3002, 1362-3095. DOI: 10.1080/09553008814551601.

[13] J. Hite and J. Mattingly, "Bayesian Metropolis methods for source localization in an urban environment," *Radiation Physics and Chemistry*, IRRMA-10, vol. 155, pp. 271–274, Feb. 1, 2019, ISSN: 0969-806X. DOI: 10.1016/j.radphyschem.2018.06.024.

[14] G. Kim, S. Jeong, J. Kim, K. T. Lim, H. Chung, and M. Kim, "Bayesian approach for inferring two-dimensional location of a radioactive material using distributed detectors," *Radiation Physics and Chemistry*, vol. 227, p. 112 389, Feb. 1, 2025, ISSN: 0969-806X. DOI: 10.1016/j.radphyschem.2024.112389.

[15] A. Bukartas, J. Wallin, R. Finck, and C. Rääf, "Bayesian algorithm to estimate position and activity of an orphan gamma source utilizing multiple detectors in a mobile gamma spectrometry system," *PLOS ONE*, vol. 16, no. 1, e0245440, Jan. 22, 2021, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0245440.

[16] A. Dvornik, R. Finck, and C. Rääf, "Enhancing Bayesian methods for radioactive source localization: A parameter study, prior construction and signal smoothing," *Journal of Environmental Radioactivity*, vol. 295, p. 107 960, Apr. 1, 2026, ISSN: 0265-931X. DOI: 10.1016/j.jenvrad.2026.107960.

[17] J. Nilsson, *Mobile Gamma Spectrometry: Development and Optimisation of Methods for Locating and Mapping Lost Radioactive Sources*. Malmö: Medical Radiation Physics, Malmö, Lund University, 2016, 66 pp., ISBN: 978-91-7619-317-4.

[18] J. S. Lilley, *Nuclear Physics: Principles and Applications* (The Manchester Physics Series). Chichester ; New York: J. Wiley, 2001, 393 pp., ISBN: 978-0-471-97936-4 978-0-471-97935-7.

[19] M. Isaksson and C. L. Rääf, *Environmental Radioactivity and Emergency Preparedness* (Series in Medical Physics and Biomedical Engineering). Boca Raton, FL: CRC Press, Taylor & Francis Group, 2017, 1 p., ISBN: 978-1-315-37287-7. DOI: 10.1201/9781315372877.

[20] UNSCEAR, "Sources, Effects and Risks of Ionizing Radiation," UNSCEAR 2008 Report, 2008, Volume 1. Annex B. Exposures of the public and workers from various sources of radiation.

[21] M. Shifotoka. "Six Global Trends in Nuclear Power You Should Know," Accessed: Mar. 30, 2026. [Online]. Available: http://www.iaea.org/newscenter/news/six-global-trends-in-nuclear-power-you-should-know.

[22] Finansdepartementet, "Long-term investments in nuclear power for Swedish electricity supply," Finansdepartementet (Ministry of Finance in Sweden), Press release, Oct. 28, 2025.

[23] Q.-H. Hu, J.-Q. Weng, and J.-S. Wang, "Sources of anthropogenic radionuclides in the environment: A review," *Journal of Environmental Radioactivity*, Selected Papers from the ICOBTE (International Conference on Biogeochemistry of Trace Elements), Beijing, China, July 2007, vol. 101, no. 6, pp. 426–437, Jun. 1, 2010, ISSN: 0265-931X. DOI: 10.1016/j.jenvrad.2008.08.004.

[24] K. E. Stenström and S. Mattsson, "Spatial and temporal variations of 14C in Fucus spp. in Swedish coastal waters," *Journal of Environmental Radioactivity*, vol. 242, p. 106 794, Feb. 2022, ISSN: 0265931X. DOI: 10.1016/j.jenvrad.2021.106794.

[25] E. Brewitz et al., "Kartläggning av omhändertagande av radioaktivt avfall från icke-kärntekniska verksamheter," Swedish Radiation Safety Authority (SSM), SSM2024-3915, 2024.

[26] S. Seltzer and J. Hubbell, *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients, NIST Standard Reference Database 126*, National Institute of Standards and Technology, 2004. DOI: 10.18434/T4D01F.

[27] A. Bukartas, "Assessment of mobile radiometry data in radiological emergencies using Bayesian statistical methods," thesis/doccomp, Lund University, 2021, ISBN: 978-91-8021-097-3.

[28] R. Finck, A. Bukartas, M. Jönsson, and C. Rääf, "Maximum detection distances for gamma emitting point sources in mobile gamma spectrometry," *Applied Radiation and Isotopes*, vol. 184, p. 110 195, Jun. 1, 2022, ISSN: 0969-8043. DOI: 10.1016/j.apradiso.2022.110195.

[29] A. Sitek and A. M. Celler, "Limitations of Poisson statistics in describing radioactive decay," *Physica Medica*, vol. 31, no. 8, pp. 1105–1107, Dec. 1, 2015, ISSN: 1120-1797. DOI: 10.1016/j.ejmp.2015.08.015.

[30] A. Johnson, Q. Ott, and M. Dogucu, *Bayes Rules! An Introduction to Applied Bayesian Modeling*. Chapman & Hall/CRC.

[31] D. McNeish, "On Using Bayesian Methods to Address Small Sample Problems," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 23, no. 5, pp. 750–773, Sep. 2, 2016, ISSN: 1070-5511. DOI: 10.1080/10705511.2016.1186549.

[32] O. Abril-Pla et al., "PyMC: A modern, and comprehensive probabilistic programming framework in Python," *PeerJ Computer Science*, vol. 9, e1516, Sep. 1, 2023, ISSN: 2376-5992. DOI: 10.7717/peerj-cs.1516.

[33] P. Nelissen. "PG-RAD - Primary Gamma RADiation Simulator, v0.1.0. "[Online]. Available: https://github.com/pim-n/pg-rad.

[34] C. R. Harris et al., "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 17, 2020, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-020-2649-2.

[35] P. Virtanen et al., "SciPy 1.0: Fundamental algorithms for scientific computing in Python," *Nat Methods*, vol. 17, no. 3, pp. 261–272, Mar. 2, 2020, ISSN: 1548-7091, 1548-7105. DOI: 10.1038/s41592-019-0686-2.

[36] E.-M. Arvanitou, A. Ampatzoglou, A. Chatzigeorgiou, and J. C. Carver, "Software engineering practices for scientific software development: A systematic mapping study," *Journal of Systems and Software*, vol. 172, p. 110 848, Feb. 2021, ISSN: 01641212. DOI: 10.1016/j.jss.2020.110848.

[37] Brandon Rhodes. "The Builder Pattern: A Creational Pattern from the Gang of Four book. "[Online]. Available: https://python-patterns.guide/gang-of-four/builder/.

[38] E. Gamma, Ed., *Design Patterns: Elements of Reusable Object-Oriented Software* (Addison-Wesley Professional Computing Series), 39. printing. Boston, Mass. Munich: Addison-Wesley, 1995, 395 pp., ISBN: 978-0-201-63361-0.

[39] A. Molnar, Z. Domozi, and I. Lovas, "Drone-Based Gamma Radiation Dose Distribution Survey with a Discrete Measurement Point Procedure," *Sensors*, vol. 21, no. 14, p. 4930, Jul. 20, 2021, ISSN: 1424-8220. DOI: 10.3390/s21144930.

[40] J. Corbacho, J. Baeza, and J. Caballero, "Use of a drone-based gamma-ray spectrometry system to assess point or extended radioactive sources," *Radioprotection*, vol. 59, no. 2, pp. 80–87, Apr. 2024, ISSN: 0033-8451, 1769-700X. DOI: 10.1051/radiopro/2023039.

A Generating pseudo-random roads

A.1 Describing a road as a planar curve

Let $r(s)$ be a planar curve describing a road in the xy-plane as a function of arc length s (the distance travelled along the road), where $s \in [0, L]$. Let $\kappa(s)$ be the curvature of $r(s)$

$$\kappa(s) = \frac{d\theta}{ds}$$

where θ is the heading (direction) of the road. Basically, $\kappa(s)$ tells us at arc length s whether the path is about to turn left or right, with larger magnitude of $\kappa(s)$ indicating a sharper change, and $\kappa(s) = 0$ indicating no change. A path in the xy-plane can now be fully defined by $\kappa(s)$ and L, as the heading angle is simply

$$\theta(s) = \int_0^s \kappa(u)\, du$$

and

$$x(s) = \int_0^s \cos(\theta(u))\, du\ ,\ y(s) = \int_0^s \sin(\theta(u))\, du$$

In practice, if we are going to generate a road in computer code we need to discretize this above formulation. Suppose a step size Δs, then at the i-th waypoint we have travelled a distance $s_i = i\Delta s$ and we have $N = L/\Delta s$ waypoints in total. At waypoint i, with known curvature κ_i, the next heading is

$$\theta_{i+1} = \theta_i + \kappa_i \Delta s$$

which by recursion means that

$$\theta_{i+1} = \sum_{j=0}^{i} \theta_j + \kappa_j \Delta s$$

and likewise

$$x_{i+1} = \sum_{j=0}^{i} x_j \cos(\theta_j)\Delta s\ ,\ y_{i+1} = \sum_{j=0}^{i} y_j \sin(\theta_j)\Delta s.$$

This shows that with starting conditions (x_0, y_0, θ_0) and discrete curvature field $\{\kappa_0, \kappa_1 ... \kappa_{N-1}\}$ we can construct a path of any arbitrary length L with its shape entirely determined by the curvature field.

A.2 Sampling roads made of pre-defined segments

Roads are typically made up of several types of sections. For example, a road between A and B may look like

straight, turn left, straight, turn right, straight

Sampling section lengths Suppose we want to build a road of length L out of K segments. The total number of waypoints N depends on the step size Δs:

$$N = \frac{L}{\Delta s}.$$

Let $(p_1, p_2, \ldots, p_K)$ represent the proportion of N that each prefab will be assigned, where $\sum_1^K p_i = 1$. One useful distribution here is the Dirichlet distribution, which is parametrized by a vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_K)$. The special case where all α_i, the scalar parameter α is called a *concentration parameter*. Setting the same α across the entire parameter space makes the distribution symmetric, meaning no prior assumptions are made regarding the proportion of N that will be assigned to each segment. $\alpha = 1$ leads to what is known as a flat Dirichlet distribution, whereas higher values lead to more dense and evenly distributed $(p_1, p_2, \ldots, p_K)$. On the other hand, keeping $\alpha \leq 1$ gives a sparser distribution which can lead to larger variance in apportioned number of waypoints to $(p_1, p_2, \ldots, p_K)$.

Suppose we draw our samples for proportion of length from the Dirichlet distribution

$$(p_1, p_2, \ldots, p_K) \sim \text{Dirichlet}(\alpha_1, \alpha_2, \ldots, \alpha_K)$$

with $\alpha_0 = \sum_{i=1}^K \alpha_i$, the mean and variance of the Dirichlet distribution are then

$$\text{E}[p_i] = \frac{\alpha_i}{\alpha_0},\ \text{Var}[p_i] = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}.$$

If α is a scalar, then $\alpha_0 = K\alpha$ and the above simplifies to

$$\text{E}[p_i] = \frac{\alpha}{K\alpha} = \frac{1}{K},\ \text{Var}[p_i] = \frac{\alpha(K\alpha - \alpha)}{(K\alpha)^2(K\alpha + 1)}.$$

We see that $\text{Var}[p_i] \propto \frac{1}{\alpha}$ meaning that the variance reduces with increasing α. We can simply scale the proportions

$$(N \cdot p_1, N \cdot p_2, \ldots, N \cdot p_K)$$

to get the randomly assigned number of waypoints for each prefab. We now have a distribution which can give randomly assigned lengths to a given list of prefabs, with a parameter to control the degree of randomness. With a large concentration parameter α, the distribution of lengths will be more uniform, with each prefab getting $N \cdot \text{E}[p_i] = \frac{N}{K}$ waypoints assigned to it. Likewise, keeping α low increases variance and allows for a more random assignment of proportions of waypoints to each prefab segment.

Sampling turn angles Suppose a turn of a pre-defined arc length l made of N/K waypoints. If one wants to create a random angle, one has to keep in mind that the minimum radius R_{min} depends on the speed of the vehicle and the weather conditions:

$$R_{\text{min,vehicle}} = \frac{v^2}{g\mu},$$

where

- v is the velocity of the vehicle in m/s,
- g is the gravitational acceleration (about 9.8 m/s^2), and

- μ is the friction coefficient (about 0.7 for dry asphalt).

A regular turn (not a U-turn or roundabout) should also have an lower and upper limit on the angle, say, 30 degrees to 90 degrees for a conservative estimate. In terms of radii, it becomes

$$R_{\text{min}} = \max\left(R_{\text{min,vehicle}}, \frac{l}{\pi/2}\right), \; R_{\text{max}} = \frac{l}{\pi/6}.$$

We then sample

$$R \sim \text{Uniform}\left(R_{\text{min}}, R_{\text{max\_angle}}\right)$$

and obtain a random radius for a turn of arc length l with limits to ensure the radius is large enough given the velocity of the vehicle. Finally, the curvature profile is related to the radius by

$$\kappa = \frac{1}{R}$$

which means that the curvature profile of a turn is simply a vector $\kappa = (1/R, \ldots, 1/R)$ with a length of N/K waypoints.

B Walkthrough: writing a config file

The config file must be a YAML file. YAML is a serialization language that works with key-value pairs, but in a syntax more readable than some other alternatives. The remainder of this chapter will explain the different required and optionals keys, what they represent, and allowed values.

Required keys

Simulation options

The first step is to name the simulation, and define the speed of the vehicle (assumed constant) and acquisition time.

Landscape name The `name` is a string, which may include spaces, numbers and special characters.
Examples:

```
name: test_landscape
```

```
name: Test landscape 1
```

Acquisition time The acquisition time of the detector in seconds.
Example:

```
acquisition_time: 1
```

Note: All units in the config file must be specified in SI units, e.g. meters and seconds, unless the key contains a unit itself (e.g. `activity_MBq` means activity in MegaBequerels).

Vehicle speed The `speed` of the vehicle in m/s. Currently, the vehicle speed must be assumed constant. An example could be

```
speed: 13.89 # this is approximately 50 km/h
```

Note: The text after the # signifies a comment. PG-RAD will ignore this, but it can be helpful for yourself to write notes.

Path

The `path` keyword is used to create a path for the detector to travel along. There are two ways to specify a path; from experimental data or by specifying a procedural path.

Path - Experimental data Currently the only supported coordinate format is the RT90 (East, North) coordinate system. If you have experimental data in CSV format with columns for these coordinates, then you can load that path into PG-RAD as follows:

```
path:
  file: path/to/experimental_data.csv
  east_col_name: East
  north_col_name: North
```

Path - Procedural Alternatively, you can let PG-RAD generate a path for you. A procedural path can be specified with at least two subkeys: length and segments.
Currently supported segments are: `straight`, `turn_left` and `turn_right`, and are provided in a list under the segments subkey as follows

```
path:
  segments:
    - straight
    - turn_left
    - straight
```

The length must also be specified, using the `length` subkey. It can be specified in two ways: a list with the same length as the segments list

```
path:
  segments:
    - straight
    - turn_left
    - straight
  length:
    - 500
    - 250
    - 500
```

which will assign that length (meters) to each segment. Alternatively, a single number can be passed:

```
path:
  segments:
    - straight
    - turn_left
    - straight
  length: 1250
```

Setting the length for the total path will cause PG-RAD to *randomly assign* portions of the total length to each segment.
Finally, there is also an option to specify the turn angle in degrees:

```yaml
path:
  segments:
    - straight
    - turn_left: 90
    - straight
  length: 1250
```

Like with the lengths, if a turn segment has no angle specified, a random one (within predefined limits) will be taken.
Note: Letting PG-RAD randomly assign lengths and angles can cause (expected) issues. That is because of physics restrictions. If the combination of length, angle (radius) and velocity of the vehicle is such that the centrifugal force makes it impossible to take this turn, PG-RAD will raise an error. To fix it, you can 1) reduce the speed; 2) define a smaller angle for the turn; or 3) assign more length to the turn segment. For more information about how procedural roads are generated, including the random sampling of lengths and angles, see Appendix A.

Sources

Currently, the only type of source supported is an isotropic point source. However, an arbitrary number of point sources can be added to the landscape. Point sources can be added under the `sources` key, where the **subkey is the name** of the source:

```yaml
sources:
  my_source: ...
```

the source name should not contain spaces or special characters other than `_` or `-`. There are three required subkeys under `sources.my_source`, which are: `activity_MBq`, `isotope` and `position`.

Source activity The source activity is in MegaBequerels and must be a strictly positive number:

```yaml
sources:
  my_source:
    activity_MBq: 100
```

Source isotope The isotope for the point source. This must be a string, following the naming convention of the symbol followed by the number of nucleons, e.g. `Cs137`:

```yaml
sources:
  my_source:
    activity_MBq: 100
    isotope: Cs137
```

Note: Currently the following isotopes are supported: `Cs137`, `Co60` and `Cs134`.

Source position There are two ways to specify the source position. Either with absolute (x,y,z) coordinates

```yaml
sources:
  my_source:
    activity_MBq: 100
    isotope: Cs137
    position: [0, 0, 0]
```

or relative to the path, using the subkeys `along_path`, `dist_from_path` and `side`

```yaml
sources:
  my_source:
    activity_MBq: 100
    isotope: Cs137
    position:
      along_path: 100
      dist_from_path: 50
      side: left
```

Note that `side` is relative to the direction of travel. The path will by default start at

$$(x, y) = (0, 0)$$

and initial heading is parallel to the x-axis.

Detector

The final required key is the `detector`. Currently, custom detectors are not yet supported and you must choose from a list of existing detectors:

```yaml
detector: LU_HPGe_90
```

Optional keys

The following subkeys are optional and should be put under the `options` key.

```yaml
options:
  air_density_kg_per_m3: 1.243
  seed: 1234
  bkg_cps: 0
```

where `bkg_cps` is the mean background count.

C Inputs

C.1 Source separation

```yaml
name: source separation test
speed: 8.33
acquisition_time: 1

path:
  length:
    - 100
    - 2000
  segments:
    - turn_left: 30
    - straight

sources:
  s1:
    activity_MBq: 300
    isotope: Cs137
    gamma_energy_keV: 662
    position:
      along_path: {a}
      dist_to_path: 32
      side: left
```

```
  s2:
    activity_MBq: 300
    isotope: Cs137
    gamma_energy_keV: 662
    position:
      along_path: {b}
      dist_to_path: 32
      side: left

detector: LU_HPGe_90

options:
    bkg_cps: 0
```

Listing 2: The input for the source separation. The distance in terms of FWHM was translated to absolute meters d, after which $\{a\} = L/2 - d/2$ and $\{b\} = L/2 + d/2$, where L is the length of the path.