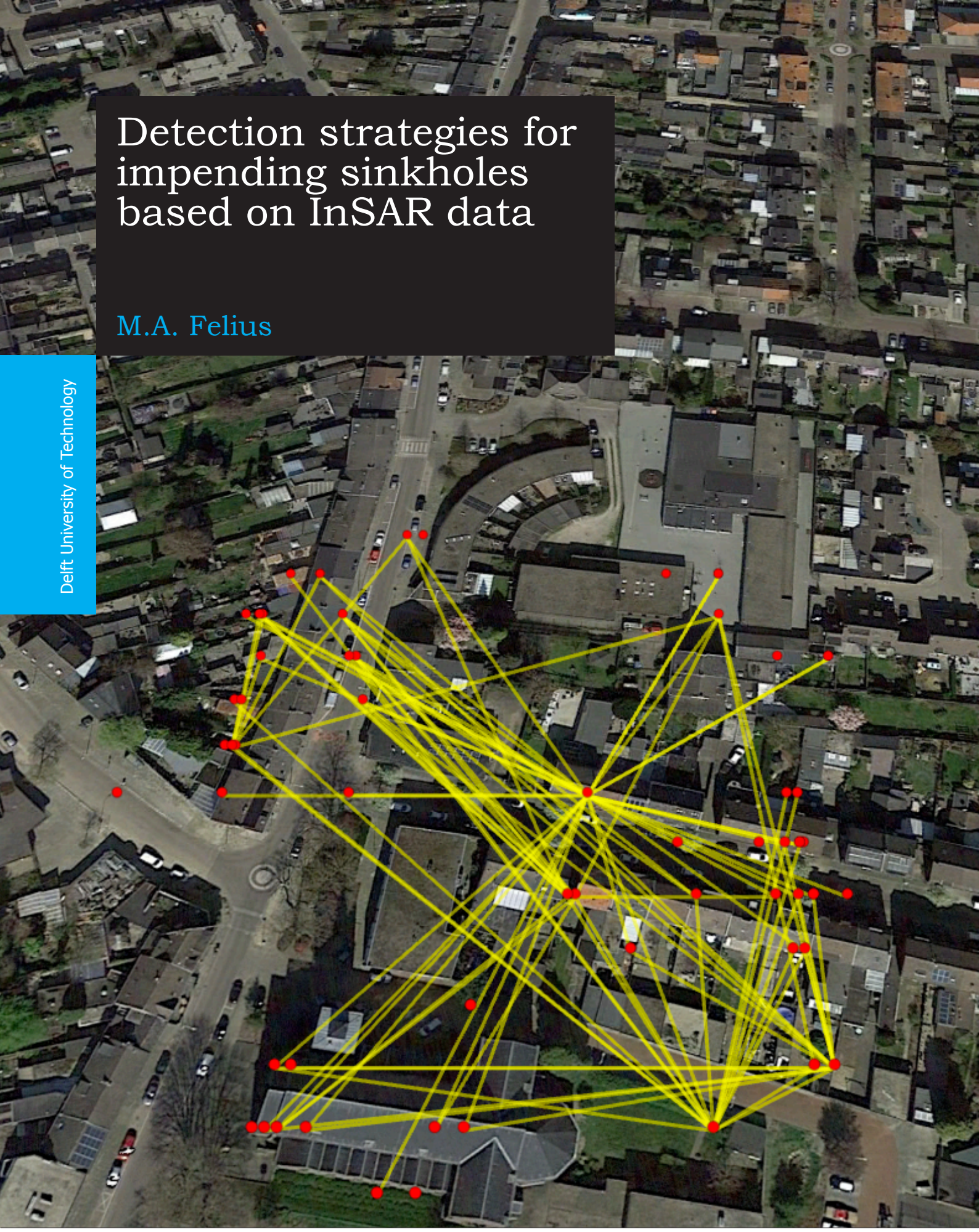
Detection strategies for impending sinkholes based on InSAR data

M.A. Felius

Delft University of Technology



TU Delft Delft University of Technology

Challenge the future

[This page was intentionally left blank.]

Detection strategies for impending sinkholes based on InSAR data

by

M.A. Felius

in partial fulfillment of the requirements for the degree of

Master of Science
in Applied Earth Sciences

at the Delft University of Technology,
to be defended publicly on Thursday September 9, 2021 at 16:00.

Supervisor:	Prof. dr. ir. R. F. Hanssen,	TU Delft, Geoscience and Remote Sensing (CiTG)
Thesis committee:	Dr. ir. F. J. van Leijen,	TU Delft, Geoscience and Remote Sensing (CiTG)
	Dr. ir. D. J. M. Ngan-Tillard	TU Delft, Geo-engineering (CiTG)

TU Delft Delft University of Technology

Preface

In the beginning, it was a challenge to get started on this thesis. I had no idea what it incorporates and how to conduct proper research. Also, working from home (due to Corona) was not an ideal start. Gradually, I learned and got a better feeling of what was expected from me. Still, I found writing very difficult, and it caused much stress in the beginning. I also managed to overcome this, and I feel much more confident writing now.

I think the field of Remote Sensing will experience rapid growth in the upcoming years. This growth will be facilitated by the increasing amount of Earth observation satellites, generating exponentially more data. With new mega-constellations, new possibilities will arise, and limits will be lifted. This study might also benefit from this trend. One limitation is the repeat-frequency of the satellite. Sinkholes are minute events; more coverage might lead to more and better detection of impending sinkholes.

I want to thank my daily supervisors, Ramon Hanssen and Freek van Leijen, for their guidance. It was a massive project to complete. I'm also very grateful for the opportunities my supervisor gave me to present my work. The presentation moments were at first daunting, but gradually I got better and more confident. I'm very proud of my last presentation at Fringe 2021. I want to thank the people from the Caroline Huddle group. These people are Wietske Brouwer, Simon van Diepen, Philip Conroy, Marc Bruna, and John Zouros. They provided a foundation for posing any questions related to the thesis.

I also want to give a big thanks to my family. My dad, Johan, my mother Marjan, my brother, Joris, and especially my little sister, Lotte, for helping me and always providing a listening ear for my frustrations.

I want to thank Skygeo for providing the data on which I could apply the developed methodology.

Max Felius
Delft, August 2021

Abstract

Undiscovered underground cavities might exist in the subsurface. A catastrophic ground failure event follows when such a cavity starts to migrate upwards and finally intersects with the surface, resulting in a sinkhole. Catastrophic collapse events are usually preceded by precursory subsidence. An upward migrating cavity causes the development of a (subsidence) trough at the surface. The trough deepens as the cavity nears the surface, changing its surface expression. With the technique called InSAR, displacements over a large area can be measured. Literature shows that precursory subsidence is measurable using InSAR. However, automatically detecting impending sinkholes from InSAR displacements has not yet been researched in depth. This study shows that the developed novel arc-based temporal strategy can early detect an impending sinkhole. The kinematic model is adequate to model the surface expression of an impending sinkhole. The results are used to implement an artificial sinkhole based on the kinematic model into a subset to test the developed strategies. The first strategy was based on spatio-temporal characteristics of a sinkhole. This strategy can locate the subsiding area and indicate a surface expression size range. However, during this study, a second strategy was developed to identify anomalous behavior quicker and more reliable than the spatio-temporal strategy. The second strategy is arc-based and marks point measurements behaving anomalously. The result demonstrates a potential automated early warning system based on the InSAR displacement time series. The developed strategies harbor the potential of monitoring for impending sinkholes on a large scale. This study is anticipated to be a starting point for more development in early warning systems for impending sinkholes. Future research could entail verifying the strategies in regions with collapsed sinkholes.

Contents

1

Introduction

1.1. Motivation

Detecting impending sinkholes is proved to work using InSAR in a retrospective fashion, but detecting sinkholes a priori is still complicated. This study focuses on new methodologies built upon existing InSAR measurements to detect sinkholes before the catastrophic collapse event.

1.2. Introduction

The Netherlands has experienced a few cases where a sudden sinkhole collapse occurred and caused significant damage. A notorious example is the sinkhole occurrence at shopping mall 't Loon in Heerlen. In an underground parking lot, a subsidence sinkhole appeared in December 2011 (Chang & Hanssen, 2014). More recently, in July 2020, a sudden collapse sinkhole formed in the city of Kerkrade. Here, a collapse sinkhole of a few meters appeared in the middle of the street. The area is infamous to host abandoned and forgotten mining shafts, and this particular sinkhole has been attributed to such an abandoned and forgotten mining shaft (Redactie Algemeen Nederlands Persbureau, 2020).

Detecting impending sinkholes is rather complex using the methods adopted from other fields of study. The features involved are usually small for the potential area affected (Intrieri et al., 2015). Moreover, the number of research papers on sinkholes is limited when compared to other ground deformation phenomena such as landslides (Gutiérrez et al., 2019).

Gutiérrez et al. (2019) categorize the monitoring methods used on sinkholes in three distinct groups; (1) subsurface methods, (2) ground-based methods, and (3) remote sensing methods.

Subsurface methods have been widely applied for investigating sinkholes. Successful methods are seismic reflection, refraction, and ground-penetrating radar (GPR). In recent years, GPR was used the most since it is favorable in detecting localized subsurface cavities (Intrieri et al., 2015).

Ground-based methods involve Robotic total stations (RTS), GPS, terrestrial laser scanner (TLS), and more. These methods all produce data with high precision and accuracy. However, for large areas, the spatial sampling needs to be very dense and is thus not feasible (Gutiérrez et al., 2019; Intrieri et al., 2015).

Remote sensing methods collect data from airborne or space-borne platforms. A standard method used is the Interferometric Synthetic Aperture Radar (InSAR) and is used for detecting precursory subsidence preceding collapse sinkholes (Intrieri et al., 2015; Jones & Blom, 2014; Malinowska et al., 2019; Nof et al., 2013). The areas containing new impending sinkholes tend to be large (> 1km). Currently, InSAR is the only suitable method covering large areas (Gutiérrez et al., 2019).

1.3. The rise of geodetic big data

Many new radar satellite missions were launched in recent years, therefore increasing the amount of data available. The Sentinel-1 missions even provide data free of charge. These developments provide for an abundant and continuous data source for developing wide-scale deformation time series.

Satellite-based radar measurements provide displacement estimates over a large area. The displacement estimates are based on a technique called Interferometric Synthetic Aperture Radar (In-

SAR) (Bamler & Hartl, 1998; Hanssen, 2001). The concept of InSAR uses at least two radar images and looks at the phase differences between those two images. From these phase differences, the relative displacement can be computed to provide the displacement estimate. Taking a set (or stack) of images and computing the relative displacements w.r.t. the first image (i.e., master image) gives a displacement time series. However, not every pixel can be linked to a decent scatterer on Earth that acts consistently through time. Scatterers who are not consistent enough through time are labeled incoherent scatterers. For example, vegetation is infamous for producing incoherent scatterers. A technique for finding coherent point scatterers in the stack of images is called Persistent Scatterer Interferometry (PSI) (Ferretti et al., 2001; Hanssen, 2001; Kampes, 2006; van Leijen, 2014).

Post facto studies have shown that precursory sinkhole subsidence is measurable using InSAR (Intrieri et al., 2015; Jones & Blom, 2014; Kim et al., 2016; Nof et al., 2013). Creating coherent InSAR images over an extended period can be challenging since individual scatterers can exhibit high phase variability. PSI is the technique that selects suitable scatterers with low phase variability over an extended period. Since PSI uses a filtering scheme, some points will get discarded. The resulting suitable scatterers tend to be heterogeneously distributed, which largely depends on the surface context. In sparse areas, it is more difficult to a priori detect precursory subsidence. However, Chang & Hanssen (2014) and Malinowska et al. (2019) show that with the reduction of available scatterers, it is still possible to detect impending sinkholes.

1.4. Different early warning systems based on InSAR Data

Multiple *ex post facto* studies have been applied on locations where a catastrophic sinkhole collapse occurred (Chang & Hanssen, 2014; Intrieri et al., 2015; Jones & Blom, 2014; Malinowska et al., 2019; Nof et al., 2019). Using InSAR, the deformation history of the area is investigated to identify potential precursory subsidence. The ultimate goal is to develop an early warning system based on InSAR that automatically detects and identifies such precursory deformation.

The wider Dead Sea area experiences a lot of catastrophic sinkhole collapse events. Over ~6000 collapse sinkholes have been mapped along the Dead Sea shores since 1980 (Nof et al., 2019), and this type of natural hazard affects the local agriculture, industry, tourism, infrastructure, and daily life in the region. Nof et al. (2019) implements a semi-automatic near-real-time early warning system. The system immediately creates interferograms when the data is available. Then, a manual inspection is needed to look for the spatial sinkhole pattern. This spatial early warning system, or spatial method, appears to be successful in the Dead Sea area. As also stated in Nof et al. (2019), the arid and sparsely populated Dead Sea area is almost ideal for monitoring impending sinkholes using InSAR. Furthermore, it suggests that insights from this operation will be beneficial for early warning systems in other, less ideal locations.

The spatial method becomes harder to implement when using the PSI technique. Chang & Hanssen (2014) and Malinowska et al. (2019) showed that within the displacement time series of a single point, an impending sinkhole signal could be detected. This temporal method looks for acceleration events indicative of an impending sinkhole. Malinowska et al. (2019) states that the results obtained in this paper should be considered supportive evidence for a potential operational application of PSI as a future sinkhole warning methodology.

Combining the spatial and temporal methods provide a new approach for a potential early warning system, the spatio-temporal method. Kratzsch (1983) and Reddish & Whittaker (1989) implement a functional model that approximates the sinkhole subsidence pattern. With the spatio-temporal method, a sinkhole pattern is assumed and subsequently searched for in the PSI dataset.

1.4.1. Sinkholes in Limburg

Limburg has a rich history of surface, near-surface, and underground mining. The first mining activities date back to the 12th century (Denys, 2019). The last mine in Limburg was closed in 1974 (Denys, 2019). However, the last mine in Germany (just across the border) was closed in 1994, keeping the dewatering measures active in the Limburg mines (Bruna, 2020). Since the closure of the last mine in Germany, the dewatering has stopped. The groundwater has been rising steadily ever since. Ground heave has been measured in Limburg and it is connected the ongoing rise of mine water.

However, Heitfeld et al. (2016) state that the risks from shafts or near-surface mining exists independent of the rising mine water. The explanation for the existence of sinkholes is related to the

presence of abandoned mine shafts which were not backfilled but collapsed. It was common practice in the past to let the roof collapse into the mined voids. Because most of these collapsed corridors are from near-surface mining activity, residual voids have to be expected close to the ground surface. Often, the layers above these voids are not thick enough to establish a stable vault. These are the situations where sinkholes might develop (Heitfeld et al., 2016).

Due to the increased risk of subsidence or sinkhole formation, the province of Limburg started a pilot study to gather suggestions for designing and establishing a Monitoring and Risk Signaling System (MRSS) for surface motion. One of the objectives of this thesis is to design a methodology to provide monitoring and signaling system for an impending calamity related to surface deformation.

1.5. Research Aim

This study investigates the feasibility of an early warning system based on Persistent Scatterer InSAR for impending sinkholes. In this context, *impending* refers to the lead time between the moment that surface subsidence becomes observable and the moment of collapse. The methodology to obtain the (PSI) estimates is based on satellite-based radar data (Hanssen, 2001; van Leijen, 2014). The early warning system should have the potential to be fully automated. Ideally, the system should have the possibility of providing push messages containing critical areas. However, before we can implement a system for detecting impending sinkholes, we need to take a closer look at the surface expression of a sinkhole, identify the limitations of the available data and define a working detecting strategy. Therefore, the main research question is:

How can an early detection of impending sinkholes be enabled using satellite-based SAR measurements?

To answer the main research question, we first investigate the relevant ground movement patterns and their characteristics. Then, we look at the inherent characteristics of a typical dataset (e.g., spatial distribution), and finally, we implement two strategies for spotting impending sinkholes. We translate these steps into the following sub-questions:

- How does a subsurface cavity manifests itself onto the surface, and how can we model the manifestation?
- How do the spatial and temporal time series characteristics influence impending sinkhole detection?
- How can a spatio-temporal strategy based on a kinematic model improve early detection of an impending sinkhole?
- How can an arc-based temporal strategy improve the early detection of an impending sinkhole?

1.6. Limitations

The topic of early sinkhole detection has many facets. In this thesis, we mainly focus on detecting a small spatial footprint and a limited lead time within the life cycle of a sinkhole.

We obtain the displacement estimates by processing raw data from satellite radar missions. Several radar missions (i.e., satellites or sensors) are regularly used: Sentinel-1, RADARSAT-2, TerraSAR-X, and COSMO-SkyMed. For the Limburg case, we had access to Sentinel-1 and RADARSAT-2, and Skygeo provided the the displacement estimates.

The process of measuring the raw data, downloading and processing to obtain the displacement estimates, takes up roughly half a day.

Each mission can provide multiple tracks. We can characterize a track by the orbit direction (ascending or descending indicated by the heading direction, α_h) and its viewing geometry (e.g., incidence angle or θ). We keep the tracks as a separate dataset since the orbital direction and viewing geometry are essential when interpreting the data (Hanssen, 2001; van Leijen, 2014). The two missions provide a total of five tracks we could use. The specific mission limits the update frequency of the tracks. The Sentinel-1 mission has a repeat orbit of 6 days, and the RADARSAT-2 mission has a repeat orbit of 24 days. The resolution is higher with the RADARSAT-2 mission. We prefer higher point densities because,

in general, it will help detect smaller sinkholes. In this thesis, we limit ourselves to one track from one particular mission. We used ascending track 88 from Sentinel-1. The reason for choosing Sentinel-1 data is because ascending track 88 covers the area of interest, and has the shortest repeat time.

We know that displacement estimates contain errors. For example, PSI data generally contains unwrapping errors. We assumed (and thus limited ourselves) that the errors in the PSI dataset were absent or insignificant and started to build our warning system on top of it.

The spatial signature we are searching for is relatively small compared to the observed area. To give an example, Chang & Hanssen (2014) describes a sinkhole with a diameter of 8 meters. The dataset used for this discovery is roughly 12 km longitude by 11 km latitude. Consequently, we created a subset. We created the subset around the Franciscanerstraat in Kerkrade as we know it is a sinkhole prone area (Redactie Algemeen Nederlands Persbureau, 2020) and Heitfeld et al. (2016) indicated that residual voids are still present in the area. However, no sinkhole was observed during the observation time. Therefore, we implemented an artificial sinkhole signal into the dataset. The subset spanned approximately 200 meters left to right and top to bottom. The strategies we are proposing can get computationally expensive and therefore we prefer to experiment on smaller subset instead of the full dataset. We also keep in mind how the strategies should be implemented on the whole dataset.

The last limitation is the observation time of an impending sinkhole. We roughly subdivide an impending sinkhole into three phase. The first stage (starting at t_0) is when the subsurface cavity forms and migrates upward. The second phase (starting at t_1) is when the subsurface cavity approaches the surface such that the induced subsidence starts to be observable (i.e., precursory subsidence). This phase is crucial because we need to detect the impending sinkhole during this phase. Phase three (starting at t_3) is the catastrophic surface collapse and end of the sinkhole life cycle. Obviously, phase two is crucial. The difficulty of this phase is that the duration is unknown. The phase can last for a few hours to a few days or even months (Gutiérrez et al., 2019). Hence, the repeat cycle of the satellite is limiting us. Sentinel-1 provides a minimum repeat time of 6 days, and it is consequently the best option.

1.7. Outline

Chapter 2 contains a literature study containing background information on sinkholes, modern detection methods, existing early warning systems, methods for mathematically describing subsidence, and PSI. Chapter 3 comprises the methodology used in this thesis. Here, we introduce the kinematic model, a linear least-squares method for estimating kinematic model parameters, the method for implementing an artificial sinkhole, maps indicating key characteristics for sinkhole identification, and the two detection strategies. Chapter 4 will then first show the results and discussion for the impact of a subsurface cavity on the surface, the implementation of the artificial sinkhole, maps containing a priori knowledge of the dataset, and the results for the two detection strategies. Finally, chapter 5 holds the conclusion and answers to the sub-questions, and recommendations.

2

Sinkhole Mechanics, Models, Detection, and Early Warning

This chapter will show the literature study. The literature study is setup as follows. First, a geological based explanation will talk about how sinkholes are classified. For the classification, the formation of a sinkhole plays a vital role and thus will be addressed. Also, triggers and the surface expression of a sinkhole will be treated (see Section 2.1). After an in depth description on sinkholes, the next section will go in depth on the characterization of ground movements (see Section 2.2). Section 2.3 will explain the different detection methods and Section 2.4 will go in-depth on how the detection InSAR technique works. This chapter ends with a section on early detection and warning systems (see Section 2.5).

2.1. Sinkhole Formation and Classification

The formation of sinkholes is widespread in karstic landscapes (Gutiérrez et al., 2007). However, sinkholes are also observed in non-karstic landscapes implicating that the geological setting is relevant but not binding. Sinkholes are closed circular phenomena that usually cause an increasing surface subsidence rate from the rim of the depression towards the center. Sinkholes can be described with a number of different characteristics. The reason for these different characteristics is the cluster of inter-related processes involved in a developing sinkhole (Waltham et al., 2005). The sinkhole depression (surface expression) displays a wide range of cylindrical, conical, bowl, or pan-shaped morphologies. The surface expression can range up to several hundreds of meters in diameter and depths up to tens of meters (Gutiérrez et al., 2008; Williams, 2003).

The term sinkhole derives from the processes of its evolution. A closed depression forms whereby water sinks into the ground, transporting sediments alongside inducing erosion (Waltham et al., 2005). In the past, geomorphologists called this phenomenon a doline. Sinkholes were a sub-category of a doline. However, American literature has seen a more liberal use of the term sinkhole. For practical purposes, sinkholes and dolines are the same, and the term sinkhole is further used in this thesis (Waltham et al., 2005).

From all the different research papers used for this thesis, we inferred that sinkholes are widespread and pose a hazard to the built environment. This section aims to explain the current state-of-the-art sinkhole classification and how a sinkhole manifests itself at the surface.

2.1.1. Sinkhole Classification

Waltham et al. (2005) classifies sinkholes into six main categories but also states that many sinkholes fall in between the categories. Gutiérrez et al. (2008) has slightly altered the classifying scheme. In following section, the classification mentioned in Gutiérrez et al. (2008) is leading and explained.

The first sinkhole class is *solution sinkholes*. The primary process of *solution sinkhole* formation is governed by differential dissolution lowering of the ground. The surface expression of a *solutions sinkholes* usually develops imperceptibly slowly (Gutiérrez et al., 2008). Therefore, the slow solution process is generally not perceived as hazardous from an engineering perspective and will not be further elaborated in this thesis (Gutiérrez et al., 2008).

The other classes of sinkholes are mainly based on two principles; the materials affected and the subsidence mechanisms. The materials affected are cover (cohesive and non-cohesive soil), bedrock, and caprock, and the subsidence mechanisms are sagging, suffosion, and collapse. Multiple combinations of terms can be used whereby a dominant material or process is followed by a secondary one (e.g., bedrock sagging and collapse).

Bedrock and caprock sinkholes

Bedrock is the rock layer that is underlying loose deposits. In the case of sinkholes, this layer is usually hosting the subsurface cavity. The caprock is any rock overlying the bedrock. The caprock drops into the void below (Waltham et al., 2005). Fig. 2.1 shows the development of different subsidence

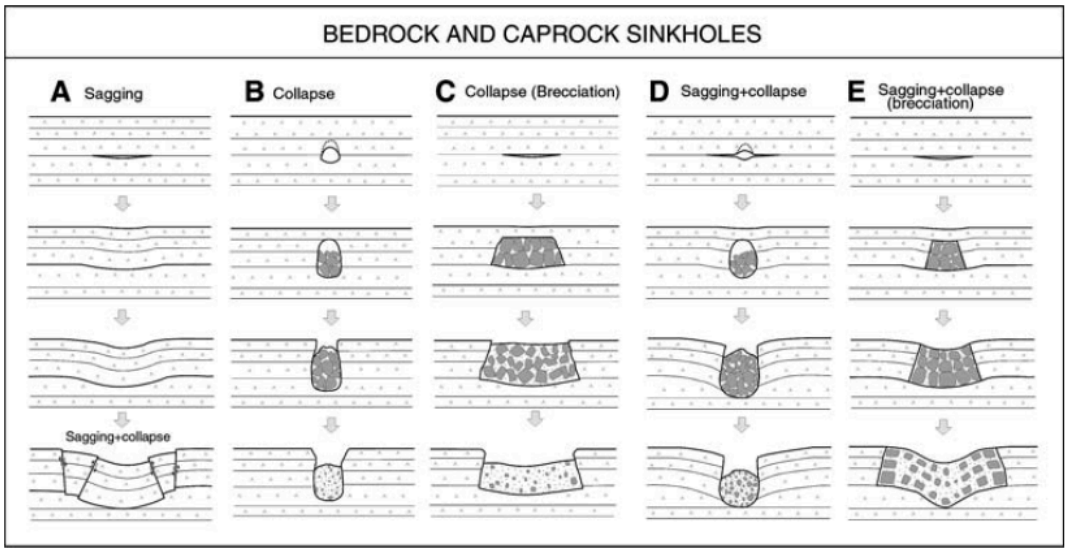


Figure 2.1: Overview of various subsidence mechanisms for bedrock and caprock sinkholes. This figure was originally published in Gutiérrez et al. (2008).

mechanisms in caprock or bedrock sinkholes. In Gutiérrez et al. (2008); Waltham et al. (2005); Williams (2003), the bedrock is usually an evaporitic layer prone to dissolution. In Fig. 2.1-A, sagging occurs where small cavities experience continuous flexure of the overlying strata so that the cavities do not necessarily develop beneath the sagging structures (Gutiérrez et al., 2008). Three factors potentially govern the development of the collapse structures in bedrock and caprock (Fig. 2.1-B to E): (1) lack of mechanical strength in the roof (Williams, 2003), (2) void roofs that are too small to initiate sagging, and (3) a vadose condition whereby the roof experiences a greater effective weight (Gutiérrez et al., 2008). The collapse of the roof produces a rubble pile on the floor. The size of the blocks depends on the type of material in the roof. As the cavity migrates upward, the pile grows, creating a collapse chimney (that is breakdown column, transtratal breccia pipe, or geological organ). These cavities may act as zones of preferential groundwater flow and accelerate rates of dissolution. The upward migrating process may cease temporarily or permanently if the cavity roof and the pile meet (Gutiérrez et al., 2008).

Cover sinkholes

The term cover incorporates both cohesive and non-cohesive soil and is used for the soil layer lying above a bedrock layer in which cavities can form. The difference between caprock and soil is that soil is not lithified. Fig. 2.2 shows three different subsidence mechanisms. These subsidence mechanisms are sagging, suffosion, and collapse. The suffosion subsidence mechanism is new and only present for cover sinkholes. The sagging mechanism in cover materials (Fig. 2.2-A) does not require the existence of cavities because continuous accommodation of the overburden material inhibits the creation of cavities. The flexure of the cover results in the development of a basin structure with centripetal dips (Gutiérrez et al., 2008). The sinkholes produced by cover sagging are generally shallow and difficult to distinguish margins; therefore, the sinkhole can be larger than expected. Although they do not form

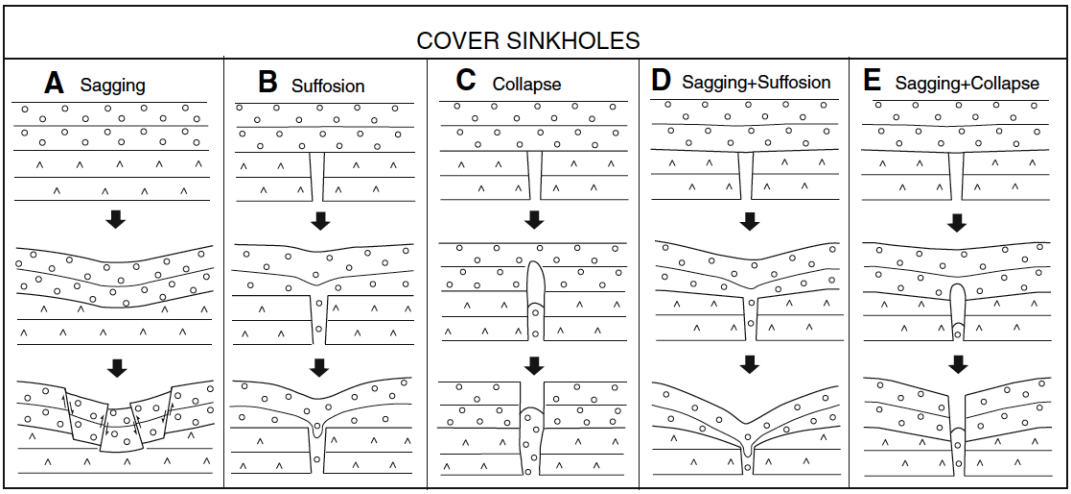


Figure 2.2: Overview of the various types of cover sinkholes. This figure was originally published in Gutiérrez et al. (2008).

a direct threat to human lives due to their slow subsidence rates, they may cause severe damage to buildings and infrastructure (Gutiérrez et al., 2008).

Cover suffosion and collapse sinkholes (Fig. 2.2-B and C) development starts at the discontinuity interface between the cover deposits and the bedrock (i.e., rockhead). Water circulation tends to concentrate at this discontinuity plane, creating cavities or pipes (conduits) of high water discharge (Gutiérrez et al., 2008). In recharge areas under vadose conditions, the enlargement of conduits by downward flow induces positive feedback. The downward flow leads to the formation of a drawdown cone around the conduit in the cover deposits affecting the local hydraulic gradient and inducing even more positive feedback (Gutiérrez et al., 2008). Unconsolidated cover deposits potentially migrate downward into the conduits by the process termed *suffosion* or *ravelling* (Waltham et al., 2005).

A few mechanisms govern the downward transport of sediments: (1) down washing of particles by percolating water; (2) cohesionless granular flow; (3) viscous gravity flow of clay-rich deposits; and (4) particle fall from the roof of the cavity (Waltham et al., 2005; Williams, 2003). Depending on the cohesion of the cover, a distinction between suffosion and collapse sinkhole is made.

Cover suffosion sinkholes generally do not form catastrophically (Fig. 2.2-B and D). The cover is a ductile or loose granular deposit. The erosional downwash of sediments and subsequently transporting out of the system undermines the cover and facilitates gradual settlement. The erosion of non-cohesive granular sand produces a funnel-shaped sinkhole whereby the slopes are governed by the angle of repose (Waltham et al., 2005). A clay-rich deposit may migrate downwards as a viscoplastic flow generating sheath folds with concentric structure (Gutiérrez et al., 2008).

Cohesive cover deposits may not be as susceptible to downwash and exhibit brittle rheology (Fig. 2.2-C and E). Therefore, arched cavities may be able to develop over conduits. The cavity can propagate progressively upward due to roof (or arch) failures. The fallen materials accumulate on the cavity floor or (with sufficient water present) get washed away. Two mechanisms can stop the upward movement, when the pile can support the roof due to bulking or when it reaches a more cohesive layer. The later mechanism causes a lateral growth whereby eventually the roof collapses (due to large arch span). This mechanism can penetrate through multiple types of layers. This upward migrating cavity will propagate upwards undetected for most of its lifespan. Depending on the cohesivity of the cover, when the cavity gets closer to the surface, sagging (subsidence) will increase gradually. The degree of sagging is dependent on the soil type (Fig. 2.2-E). Eventually, the collapse of the cavity roof may intercept the ground surface and abruptly creating a (catastrophic) collapse sinkhole (Gutiérrez et al., 2008).

Sinkhole Triggers

Specific events can trigger the creation of a subsurface cavity. Gradual dissolution can trigger sinkholes, but other factors, including humans, can also induce sinkholes to form. British Geological Survey (2020) lists a few key triggers, these are:

- Heavy rain or surface flooding can initiate the collapse of normally stable cavities, especially those

developed within superficial deposits.

- Leaking drainage pipes, burst water mains, irrigation, or even the act of emptying a swimming pool are all documented examples of sinkhole triggers.
- Construction and development are also potential triggers. Modifying surface drainage or altering loads imposed on the ground without adequate support can cause sinkholes to develop.
- Drought or groundwater abstraction can cause sinkholes by changing the level of the water table. A declining water table removes the buoyant support water provides to a cavity. Draining these cavities can cause them to collapse.
- Mining can cause sinkholes by dewatering and lowering the water table or intercepting clay-filled voids that subsequently collapse.
- Gypsum, limestone, and salt layers are susceptible to dissolution, which can cause underground cavities to form.

Knowing the potential triggers for a specific area will make it more insightful what the hazards associated with sinkholes are. These hazards can then be mitigated by appropriate planning, good site investigation, appropriate design and proper maintenance of infrastructure such as drains and services (British Geological Survey, 2020). Knowledge of the existence of specific triggers in an area will also help with identifying surface characteristics or expressions of a sinkhole.

2.1.2. Sinkhole Surface Expression

Collapse sinkholes have devastating consequences and potentially a significant impact when reaching the surface. An extreme example is a collapse sinkhole in Guatemala. On February 22, 2007, a collapse sinkhole appeared at the intersection of two streets. The sinkhole was 30 m in diameter and 60 m deep. The collapse destroyed houses and even costed people their lives (Hermosilla, 2012). The sinkhole in Guatemala is an extreme example, but many other examples of sinkhole formation showed a significant hazard to the built environment (Chang & Hanssen, 2014; Gutiérrez et al., 2007; Intrieri et al., 2015; Malinowska et al., 2019; Nof et al., 2013).

The sinkhole visible at the surface is only a small sign of the extent of the actual cavity underground (Augarde et al., 2003; Hermosilla, 2012). The surface expression is the (sub)circular shape of the subsidence bowl visible at the surface due to the influence of an upward migration cavity (impending sinkhole). The surface expression is usually the best predictor since other methods for detecting impending sinkholes have a high cost or low spatial resolution. High spatial resolution is important because sinkhole prone areas tend to be very large relative to the sinkhole expression (Intrieri et al., 2015).

Some sinkholes have a minimal surface expression since the cover material is very strong (i.e., cohesive or caprock/bedrock) and are therefore hard to detect (e.g.,Fig 2.2-C and Fig. 2.1-B). When looking from an observational perspective, the detected surface expression can describe by two classifications. These are the *cover-collapse* and *cover-subsidence* sinkholes (Hanssen et al., 2020). Cover sagging, suffosion, and collapse sinkholes are classified as subsidence sinkholes in Waltham et al. (2005). Therefore, from an observational viewpoint, the distinction is made that a sinkhole forms as a *cover-subsidence* sinkhole and potentially develops into a *cover-collapse* sinkhole.

The process of this development can be separated into three phases. The first phase is cavity creation (trigger) and upward migration. The depth of the cavity is such that deformations at the surface are negligible. The second phase starts when the cavity approaches the surface, and interactions between the surface and the cavity start to arise. This means that for each cavity position in the subsurface, the surface will deform accordingly. In the second stage, the cavity migration can (temporarily) slow down or stop due to geological setting or the phreatic water level. The third and last phase is the (catastrophic) collapse whereby the cavity intersects with the surface creating a sinkhole (Hanssen et al., 2020).

The second phase is of most interest to this thesis. During this phase, the cavity exerts influence on the surface, which we want to detect. These morphological features are mainly controlled by the type of material the sinkhole is forming in. Weak or soft materials are characterized by a near-surface strain distributed across many small fractures. A gradually and smoothly developing subsidence bowl at the

surface widens as it deepens before the collapse (if that happens). Stronger materials tend to show less subsidence before the collapse. Also, the sinkhole widens as it deepens but more in a step-wise manner. New marginal fractures form and delimit marginal blocks (Al-Halbouni et al., 2019).

The geometric relationship between a subsurface cavity and sinkhole's surface expression is not straightforward and should be treated with caution (Al-Halbouni et al., 2019). However, in a stronger material, the sinkhole's surface expression may relate to subsurface cavity geometry to a variable degree. It means that the relation between the cavity growth and the surface expression can be inferred, to a certain extent. This relationship is especially strong for shallow cavities, overlying with strong material and stable cavity growth up until the catastrophic surface collapse (Al-Halbouni et al., 2019).

The geological setting and water availability govern the speed by which a sinkhole develops (Gutiérrez et al., 2008; Luu et al., 2019).

Upward migrating cavities might encounter more cohesive layers, which (temporary) slows down the upward movement. After traversing these cohesive layers, acceleration events can occur (as observed in Chang & Hanssen (2014)). However, sometimes the upward migration completely stops due to the more cohesive layers (Gutiérrez et al., 2008). Sinkholes might also exhibit accelerated subsidence by cavities migrating to softer soils (Nof et al., 2019). Sinkhole formation also tends to increase with an increasing water supply (e.g., flooding) (Luu et al., 2019). Higher rates of water will result in a higher transport rate of sediments. Luu et al. (2019) also concludes that the floodwater level does not seem to influence sinkhole size substantially.

Modeling Sinkhole Surface Expressions

Understanding the hazard posed by impending sinkholes can be precious information. However, most of the sinkhole development happens underground, and most of the time, sinkholes are only detected after their catastrophic collapse event has occurred. Sinkhole models try to relate the surface expression to the subsurface cavity development. Nonetheless, this linking process should be treated with caution (see Al-Halbouni et al. (2019)).

The literature describes various methods to model a developing sinkhole. Shalev & Lyakhovsky (2012) describes the use of a 2-dimensional numerical viscoelastic damage rheology model. Here, ground failure concerning viscous flow, brittle fractures, or a combination is simulated by step-wise removing material. Al-Halbouni et al. (2019) describes the use of a 2-dimensional distinct element method for numerical simulations. It simulates cavity growth and relates that to the surface expression. Atzori et al. (2015) and Kim et al. (2016) describe the use of the Okada model (Okada, 1985) to compute surface displacements. The Okada model is calibrated to the local case using InSAR displacement estimates (Hanssen, 2001; van Leijen, 2014) obtained from space-borne measurements. Kratzsch (1983) and Ren et al. (1987) describe the development of the surface expression using empirically derived mathematical models. These are numerous models used to approximate and predict sinkhole behavior. Every model has its benefits and its limitations. Therefore, each case will have their own preferred model.

Models can approximate the general behavior of an impending sinkhole. In turn, these models can be calibrated to fit a specific case using obtained ground samples and ground displacement measurements. However, calibrating a model to fit a specific case might be costly to achieve. Numerical models of the interaction between cavity growth, material deformation, and overburden collapse is desirable to understand sinkhole hazard better but is a challenging task due to the involved high strains and material discontinuities (Al-Halbouni et al., 2019). Perhaps a numerical model is only viable in areas with a high frequency of sinkhole formation and relatively unified and simple soil makeup. Therefore, simpler sinkhole surface expressions (i.e., Kratzsch (1983) and Ren et al. (1987)) might be desirable in new areas, areas with low sinkhole formation frequency, or areas where both conditions are combined.

2.2. Characterization of Ground Movement

When a subsurface void is created, it impacts the stress distribution in the subsurface. The movement of the strata is often compared to the bending of a beam or slab (Eckardt, 1913). If the void gets too large, the roof collapses, filling up the void. The upward migrating process can continue until it intersects with the ground surface and ending in a catastrophic collapse event. When the void gets closer to the surface, the induced stresses will increasingly be more visible as a surface depression. The

surface depression commonly looks like a subsidence trough whereby the soil predominantly moves downward and, to a lesser degree, moves horizontally towards the center of the trough.

This section aims to quantify the effect of a subsurface void on the vertical and horizontal surface displacements. This quantification for the vertical displacements results in mathematical expressions for subsidence, slope, and curvature. The horizontal movement is quantified as horizontal displacement, linear change, and compression or extension.

There are two viewpoints to define and calculate vertical and horizontal displacement. The first viewpoint is the Eulerian viewpoint. Here, the horizontal location is fixed, and the point elevation changes at that location. The Lagrangian viewpoint follows a particular point and looks at the various forces applied to that specific point. The hypothesis is that the Lagrangian viewpoint corresponds better to a PSI point since it follows a moving reflection point. The reflection point moves in the north, east, and up direction. The three directions are projected as one vector onto the line of sight direction.

2.2.1. Components of Ground Movement

As already mentioned, the subsurface void induces vertical and horizontal displacements. These displacement directions are referred to as *components*. The overarching term *ground movement* is therefore split up into two components, vertical and horizontal displacement. The downward vertical displacement is usually referred to as subsidence. Derivative components of the subsidence are the slope (or tilt) and the curvature (Kratzsch, 1983; Ren et al., 1987). The horizontal displacement acts in all horizontal directions. The *linear change* is the change in horizontal distance between two consecutive and discrete points. The compression or extension is calculated by dividing the linear change by a particular length (e.g., 1 meter) (Kratzsch, 1983).

The creation of a subsurface void induces stress, and stress is a physical quantity that expresses internal forces of neighboring particles exert on each other. Strain is a measure of the deformation of a material—the material breaks or ruptures when the strain gets too large. The breaking or rupturing of the materials is dependent on the material properties.

A void approaching the surface causes a subsidence trough to arise at the surface. This subsidence trough has a particular slope and curvature. Also, some horizontal displacement is present during the formation of the subsidence trough.

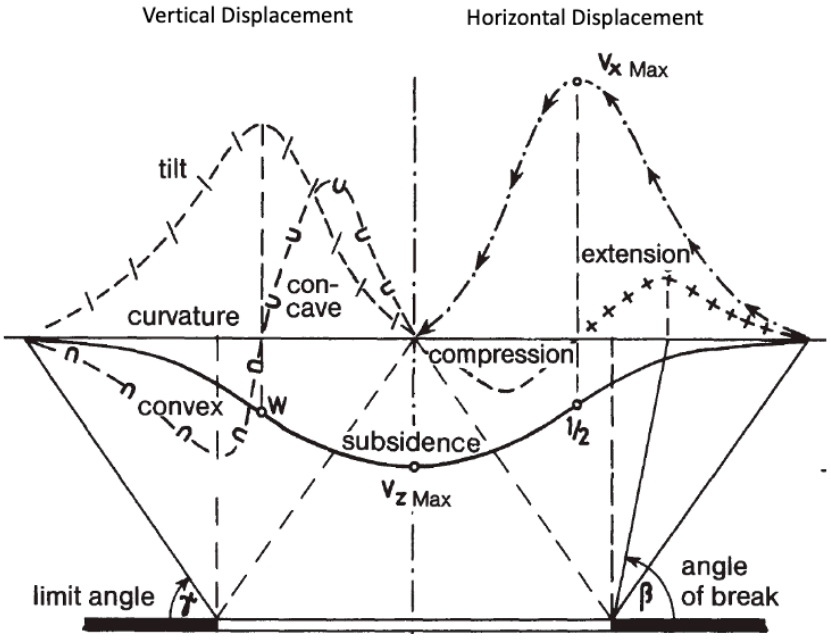


Figure 2.3: Terminology and expected displacement for different components of ground movement induced by a subsurface void. Figure from (Kratzsch, 1983).

Subsidence

The left-hand side of Fig. 2.3 shows the subsidence component and derivatives of that component. The first curve is the subsidence curve creating the subsidence trough. At the center of this subsidence trough, the maximum subsidence ($v_{z,\max}$) is located. Taking the first derivative of the subsidence trough will give the slope, slant, inclination, or tilt (v_z'). The maximum tilt is denoted as W in Fig. 2.3. Taking the derivative of the tilt gives the curvature (v_z''). The curvature is subdivided into two parts, the convex part and the concave part. The last parameter of interest for the subsidence calculation is the limit angle (γ). The limit angle shows the *limit* of the influence cone of the subsurface void on the surface.

Horizontal Displacement

The right hand side of Fig. 2.3 shows the horizontal component and its derivatives. The displacement curve resembles the curve of the tilt (Kratzsch, 1983) but is not the same. The first curve is the horizontal displacement with the maximum displacement denoted as $v_{r,max}$. The horizontal displacement acts radially towards the center of the subsidence trough. Consider that the surface consists of discrete points, the amount of horizontal displacement on each point depends on its location in the subsidence trough and results in changing point distances. The changing distance between discrete points is called the *linear change*. From the edge of the trough towards maximum horizontal displacement, extension strain ($+\epsilon$) is observed. At the point of maximum extension strain, the angle of break is defined (Fig. 2.3). The angle of break is the first location where ruptures will arise (Kratzsch, 1983). Compressive strain ($-\epsilon$) is observed from the maximum horizontal displacement towards the center of the trough.

2.2.2. Lagrangian and Eulerian Viewpoint

The two viewpoints originate from fluid mechanics. The Langrangian viewpoint of fluid mechanics extends particle mechanics and focuses on the material particles as it moves through the flow. Each of these particles is identified by its original position. The Eulerian viewpoint looks at a specific location in space and observes that location as time proceeds. With the Lagrangian viewpoint, all positions of interest are moved w.r.t. a common location in space. In contrast, with the Eulerian viewpoint, the position in space depends on the position affected (Panton, 2006). Adopting the Eulerian viewpoint loses the ability to track the history of a particle, and the attention is focused on specific points in space at various times.

Therefore, measuring and modeling techniques either measure from a Lagrangian or Eulerian viewpoint; for example, leveling measures from an Eulerian viewpoint since it is only reports vertical movement. Global Navigation Satellite System (GNSS) measures the North, East, and Up direction of a physical benchmark. Therefore identifying the origin is still possible.

Models can make use of either viewpoint. When the model is only interested in the subsidence and tries to estimate the new shape of the surface, it is referred to as geometry-driven modeling (e.g., kinematic deformation model). In contrast, when a particular physical process estimates the new surface shape, it is called physics-driven modeling (e.g., Mogi source).

With InSAR, the displacement time series is created relative to the first epoch (Hanssen, 2001; van Leijen, 2014). Therefore, the origin of a point is automatically defined, and the movement is tracked by the satellite over time. These characteristics are more in line with the Lagrangian viewpoint than the Eulerian viewpoint. However, most (geometric) models are based on Eulerian factors. The Mogi model uses the starting location of a point in space and calculates the displacement a point undergoes in three directions. Thus, the Mogi model uses the Lagrangian viewpoint to calculate ground displacements.

Lagrangian Mogi Point Source Model

The *Mogi model* is a physics-driven model based on the principle of an inflating or deflating magma chamber, which induces surface displacements (Delaney & McTigue, 1994; Dzurisin, 2006; Kiyoo, 1958). An analytical solution is obtained by introducing a spherical cavity that inflates or deflates due to hydrostatic pressure change in an elastic half-space. The equation is:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = a^3 \Delta P \frac{(1-\nu)}{\mu} \begin{bmatrix} \frac{x}{R^3} \\ \frac{y}{R^3} \\ \frac{d}{R^3} \end{bmatrix}, \tag{2.1}$$

where u,v,w are displacements of the point (x, y, z), a is the radius of the hydrostatic sphere, ΔP is the change in hydrostatic pressure in the sphere, μ is the shear modulus, the center of the cavity is at $(0, 0, -d)$, and $R = \sqrt{x^2 + y^2 + d^2}$ is the radial distance from the center of the cavity to a point on the surface. ν is Poisson's ratio of the half-space. Three assumptions are critical when using Eq. (2.1):

1. the subsurface is assumed to behave as an isotropic elastic half space (Poisson's ratio ν, Shear modulus μ),
2. the sphere's radius, a, is much smaller than the depth of the point source, d. (i.e., $a << d$).

Delaney & McTigue (1994) found a relationship between the hydrostatic pressure and the volume change of a point source. The relation is

$$\Delta V \approx \frac{\Delta P}{\mu} \pi a^3. \tag{2.2}$$

Implementing Eq. (2.2) into Eq. (2.1) gives

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \Delta V \frac{1-\nu}{\pi} \begin{bmatrix} \frac{x}{R^3} \\ \frac{y}{R^3} \\ \frac{z}{R^3} \end{bmatrix}, \tag{2.3}$$

where the cavity radius a, shear modulus μ and the hydrostatic pressure ΔP are substituted by π and ΔV (Delaney & McTigue, 1994; Dzurisin, 2006; Kiyoo, 1958).

Eulerian-based Geometric Models

Analyzing the movement of points in a trough area shows that the points move in the direction of a *focal point* in the subsurface (Kratzsch, 1983). In the case of an impending sinkhole, the focal point can also be referred to as the subsurface void or cavity. Therefore, the subsurface void at depth H creates at the surface a subsidence trough with a specific shape. The deepest point of this trough lies directly vertically above the void. The margin of the surface trough is determined by the limit angle γ (Kratzsch, 1983; Reddish & Whittaker, 1989; Ren & Li, 2008). Fig. 2.4 shows the different angles of

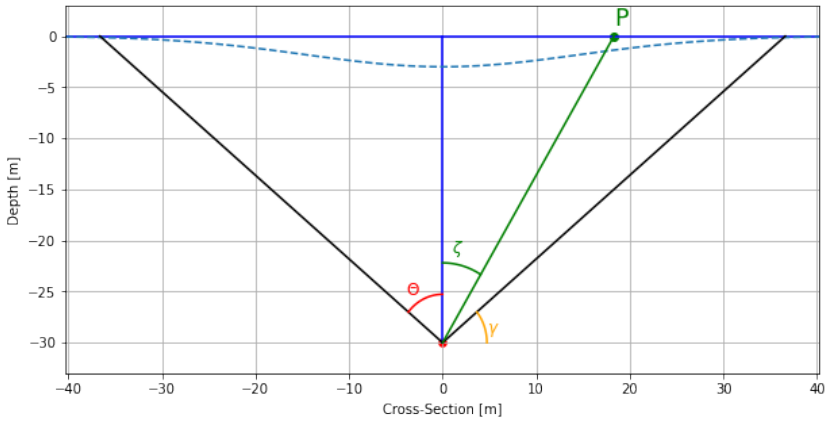


Figure 2.4: Simplified sinkhole model showing the different angles of interest. Θ is the angle of draw, ζ is the zone angle and γ is the limit angle.

interest in determining trough parameters. ζ is the zone angle that lies within the range $0 \leq \zeta \leq \Theta$. The radius of influence (R) is then computed using

$$R = \frac{H}{\tan(\gamma)} \tag{2.4}$$

and shows the extent to which the subsurface void exhibits its influence onto the surface. The limit angle is a critical parameter in subsidence analysis and prediction according to Ren & Li (2008). The

limit angle defines the extent of the effect a subsurface cavity has on the surface. Beyond this limit, the impact of subsidence is insignificant and negligible. Also, the angle depends on the strength and the composition of the overburden (Ren & Li, 2008). Therefore, taking local geological aspects into account. A consequence is that the R is the size of the surface expression and beyond this distance the subsidence is expected to be insignificant and negligible.

Selected geometric models (influence functions)
The influence function (i.e., geometric model) determines the shape of the trough. An influence function can either be derived analytically from empirical evidence or is based on assumptions. Several influence functions have been proposed this way. A benefit of influence functions is their ability to predict subsidence movements at any point above a subsurface void of any practical shape (depending on the function used). Since the influence functions rely mainly upon the radius of influence (R), they can be readily adapted to take different geological conditions into account (Reddish & Whittaker, 1989).

The most well-known and common influence function is the stochastic (Gaussian) influence function (Reddish & Whittaker, 1989; Ren et al., 1987). The function is given by

$$k_z = \frac{1}{R^2} \exp\left(-\pi \frac{r^2}{R^2}\right). \tag{2.5}$$

However, other selected functions are also proposed, such as

$$k_z = \frac{3}{\pi R^2}\left[1 - \left(\frac{r}{R}\right)^2\right]^2, \tag{2.6}$$

and

$$k_z = \frac{2}{\pi\sqrt{\pi}R}\frac{1}{r}\exp\left(-4\left(\frac{r}{R}\right)^2\right). \tag{2.7}$$

Other influence functions are described in Ren et al. (1987),

$$k_z = \cos^2\zeta, \tag{2.8}$$

and

$$k_z = 2.256\frac{1}{r}\exp(-4r^2), \tag{2.9}$$

and

$$k_z = 0.1392\exp(-0.5r^2). \tag{2.10}$$

Eq. (2.8) makes use of the ζ angle, the zone angle. The zone angle is in the range of $0 \leq \zeta \leq \theta$, whereby θ is the angle of draw. These angles are not to be confused with the limit angle γ. Fig. 2.4 shows an overview of the different angles.

2.2.3. Mathematical Definition of Ground Movement Components

The subsidence is calculated using an influence function discussed in the previous section. A multitude of influence functions were proposed by researchers (Kratzsch, 1983; Reddish & Whittaker, 1989; Ren et al., 1987). However, the stochastic (Gaussian) influence function is the most popular function for subsidence modeling (Ren & Li, 2008) and will be implemented in this thesis. The Gaussian influence function is shown in Eq. (2.5). For deriving the different ground movement components, Eq. (2.5) is adjusted to

$$k_z = \exp(-\pi \frac{r^2}{R^2}). \tag{2.11}$$

The adjusted version of the equation is further used in this thesis and is further expanded on in Chapter 3. The influence function can be scaled by multiplying it with a *scaling factor*, denoted as S. Combining both the influence function and the scaling factor provides

$$v_z = S \cdot k_z \tag{2.12}$$

where v_z is the experienced subsidence at an arbitrary location within the trough. The equation becomes

$$v_z = S \exp(-\pi \frac{r^2}{R^2}). \tag{2.13}$$

Eq. (2.13) is the equation that is scaled to provide the *approximate* subsidence at every position within the trough. The v_z is **not** the subsidence velocity. It is the subsidence and the same terminology is adopted from Kratzsch (1983).

Subsidence

The scaled influence function gives the subsidence. The tilt and curvature can be derived from the subsidence curve. There are two approaches identified to calculate the tilt and the curvature based on the subsidence curve. The first analytical approach takes the first and second-order derivatives of Eq. (2.13) to calculate the tilt and the curvature of the subsidence profile. The second approach is the numerical approximation, whereby numerical techniques are used to calculate the tilt and the curvature.

The first approach is to use Eq. (2.11) and taking the first-order derivative. The equation becomes

$$k_z' = \frac{2\pi r}{R^2} \exp(-\pi \frac{r^2}{R^2}). \tag{2.14}$$

The second derivative of the Gaussian influence function is

$$k_z'' = 2\pi \frac{R^2 - 2\pi r^2}{R^4} \exp(-\pi \frac{r^2}{R^2}), \tag{2.15}$$

and is used to determine the curvature of the profile. Taking the inverse of the curvature produces the *Radius of Curvature* ρ_z,

$$\rho_z = \frac{1}{k_z''}, \tag{2.16}$$

and is the radius of a circle that intersects the curve at a specific point. The curvature is convex in the area of the trough margin and concave in its inner zone. At the transition zone (location of maximum tilt) the curvature goes from convex to concave (Kratzsch, 1983).

The second approach is the numerical approximation, which uses numerical techniques to derive the subsidence curve's tilt and curvature. The first derivative numerical equation is written as

$$k_z' = \frac{k_z^{i-1} - k_z^i}{s}, \tag{2.17}$$

and is the backward difference equation for the first derivative whereby k_z^i is a point on the subsidence curve, and s is the distance between two consecutive points. Kratzsch (1983) provides numerical equations to calculate the concave and convex parts of the subsidence curve. The equation for the concave part of the subsidence curve is

$$k_z'' = \frac{2k_z^i - k_z^{i-1} - k_z^{i+1}}{s^2}, \tag{2.18}$$

whereby points from three locations are used. The equation for the convex part is

$$k_z'' = \frac{k_z^{i-1} - 2k_z^i + k_z^{i+1}}{s^2}. \tag{2.19}$$

These equations are based on the central difference of the second derivative (Heath, 2018). As stated earlier and in Kratzsch (1983), the curvature is convex in the area of the trough margin and concave in its inner zone.

Horizontal Displacement

One technique to directly measure subsidence is through leveling. Horizontal displacement can be determined trigonometrically with a total station, using angles and distances measured from a known point (Kratzsch, 1983). Therefore, measuring the horizontal movement is not always available. Kratzsch (1983) states that the movement of points in a trough area displays movement in the direction of a *focal point* in the subsurface. In this section, the horizontal displacement is calculated using the *focal point method*.

Focal Point Method

The focal point method assumes that there is a known depth for the cavity. The depth of the cavity is challenging to determine from just the surface subsidence pattern. However, using the relation shown in Eq. (2.4) an estimate of the depth of the cavity can be made. Also, as will be shown in the following chapters, the shape of the curve will tend to be independent of the cavity depth, and the cavity depth only has an influence on the magnitude of the horizontal displacement and its subsequent components.

Then, the equation for the focal point method becomes

$$v_r = v_z \frac{r}{H}, \tag{2.20}$$

where r is the radial distance from the center, H is the cavity depth, and v_z the subsidence at radial distance r. The geometry is shown in Fig. 2.5. Here, we see the parameters from Eq. (2.20). The

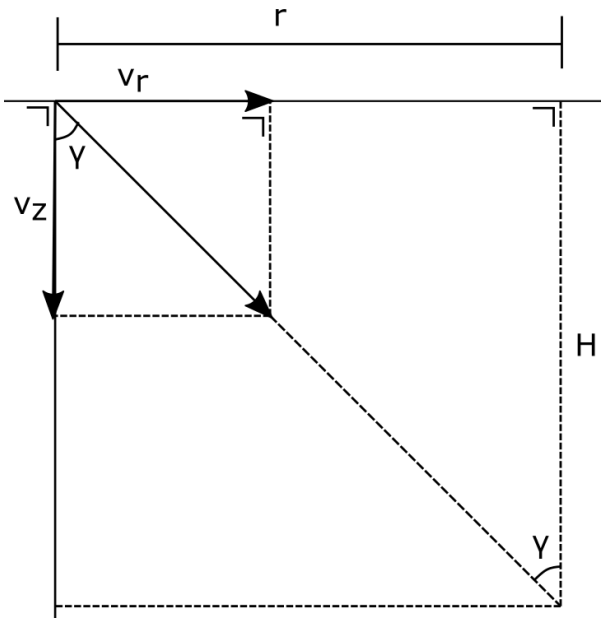


Figure 2.5: Simple cross section on the calculation of the horizontal displacement using the known subsidence based on the focal point method.

figure also shows the γ parameter which is the limit angle and used in Eq. (2.4).

The horizontal displacement is calculated using Eq. (2.20) with the parameters subsidence (v_z), radial distance (r) and the depth (H) of the cavity. Since the radial distance is taken, it will result in a figure which is (concentric) symmetric w.r.t. the center of the impending sinkhole.

Linear Change

The magnitude of the horizontal displacement differs with the radial distance towards the center. The non-uniform horizontal displacement will lead to changes in point spacing. The term for changes in point spacing is dubbed linear change, or mathematically written as Δs. Linear change is expressed as

$$\pm \Delta s = v_{x2} - v_{x1}, \tag{2.21}$$

which is the difference in the horizontal position of two consecutive points. When the linear change is negative, shortening is happening. When the linear change is positive, lengthening is taking place.

Compression and Extension
When relating the linear change to a fixed distance (s), usually 1 meter, the compressive and extensive strain ($\pm\epsilon$) is calculated:

$$\epsilon = \pm\frac{\Delta s}{s} = \pm\frac{v_{x2} - v_{x1}}{s}. \tag{2.22}$$

Compressive strain is negative, and extension strain is positive. The angle of break is calculated from the maximum value of the extension strain. The angle of break is shown in Fig. 2.3 and explained in Section 2.2.1. Table 2.1 summarizes the different ground movement components and its derivatives.

Table 2.1: Table from Kratzsch (1983) showing the different ground movement components together with their symbol and units.

Vertical Components			Horizontal Components		
Designation	Symbol	Unit	Designation	Symbol	Unit
Subsidence	v_z	m	Displacement	v_x,v_y,v_r	m
Tilt	v_z'	m/m	Linear Change	$\pm\Delta s$	m
Curvature	v_z''	1/m	Extension	$+\epsilon$	m/m
Radius of influence	ρ_z	m	Compression	$-\epsilon$	m/m

2.3. Detection Methods and Conditions

The easiest way of searching for impending sinkholes is by direct inspections in the field. Subtle cues might reveal the presence of impending sinkholes in the area, e.g., direct observation of cracks in buildings (Buchignani et al., 2007).

Other detection techniques rely on two characteristics of a sinkhole; the existence of a subsurface cavity or the detection of precursory motion .

The detection of subsurface cavities can be measured using various geophysical measurement techniques. The five most common geophysical techniques are electrical resistivity, electromagnetic conductivity (EM), ground-penetrating radar (GPR), microgravimetry, and cross-hole tomography (Gutiérrez et al., 2007). The technique that is most frequently used in recent years is GPR. GPR has a rapid acquisition of data and can identify and localize subsurface geometries and cavities (Gutiérrez et al., 2007; Intrieri et al., 2015). The disadvantage of using GPR is its limited depth, and penetration is greatly reduced by conductive materials such as clay and water.

Detection of precursory motion can be measured using geodetic techniques. Relevant geodetic (monitoring) techniques to measure surface motion are robotic total stations (RTS), global positioning systems (GPS), and terrestrial laser scanning (TLS). The monitoring technique RTS measures distances and angles with an accuracy at the millimeter level. The limitation in using this technique is that the targets should not be placed further apart than the sinkhole's diameter (Intrieri et al., 2015). RTS would be increasingly impractical for smaller sinkholes or larger areas. GPS is better suited for monitoring sinkholes. It has a high acquisition frequency and in certain configurations it can even obtain millimeter precision (Elnabwy et al., 2013; Intrieri et al., 2015; Wang et al., 2011; Zhu et al., 2014). However, the limitation of GPS is that it will only measure a small area. Covering large areas can only be achieved by placing many antennas (Intrieri et al., 2015). The TLS has been primarily used to study well-developed sinkholes. The limitations are that the presence of vegetation will influence the measurements, and in order to measure subsidence, the angle should be high between the laser scanner and the ground (Intrieri et al., 2015).

Remote sensing methods collect data from airborne or spaceborne platforms. A standard method used is Interferometric Synthetic Aperture Radar (InSAR) which can be used for detecting precursory subsidence preceding collapse sinkholes (Jones & Blom, 2014; Kim et al., 2016; Nof et al., 2013). The areas containing new impending sinkholes tend to be extensive ($>$ 1km). Currently, InSAR and PS-InSAR (Nof et al., 2019) are the only suitable methods covering large areas (Gutiérrez et al., 2019).

To successfully detect impending sinkholes, Hanssen et al. (2020) has composed four detection conditions necessary to detect impending sinkholes. Each item is necessary to observe changes in the

surface geometry. These conditions are:

1. sufficient spatial sampling (at least two measurement points close enough to the center)
2. sufficient temporal sampling (at least two measurement points within the precursory subsidence interval)
3. sufficient precision (to detect precursory subsidence)
4. sufficient representativity, where the measurements are representative of the deformation signal of interest.

These conditions were proposed in Hanssen et al. (2020) and act as guidelines for detecting impending sinkholes.

2.4. Persistent Scatterer Interferometric Synthetic Aperture Radar

Large-scale surface monitoring has had a massive influx of new applications due to the increasing availability of satellite-based data. Such a new application is the surface deformation measurements using satellite-based radar data. With radar surface deformation measurements, mm-scale deformations can be observed and, ideally, coupled with different phenomena such as sinkholes. In this chapter, the steps from satellite radar measurements to PSI displacement estimates is explained,
since this study needs to detect precursory subsidence precisely.

2.4.1. Radar

Radar stands for 'Radio Detection and Ranging' and is both a technique and a method. With radar, electromagnetic (EM) pulses are transmitted and received with either the same antenna (monostatic) or a different antenna (bistatic). By timing the time-of-flight of the radar pulse, the distance can be computed. Radar is an active system, meaning it sends out signals and does not rely on external sources (Hanssen, 2001). It can be used in all weather conditions and during day and night (Skolnik et al., 1962).

2.4.2. SAR

A (satellite) coherent radar sends out a pulse, or chirp, and waits for the response. It measures the received phase and amplitude of the signal. Synthetic Aperture Radar (SAR) is a technique that looks at stable phase behavior within the period of sending and receiving of the pulse. In this way, the azimuth resolution (the direction, heading of the satellite) will increase about three orders of magnitude (Hanssen, 2001). The obtained signal is in the format of a Single Look Complex ($a + bi$) imaginary number, but can also be denoted as a complex phasor, $P = A\exp(i\Phi)$ whereby A is the amplitude of the signal and Φ the fractional phase (Hanssen, 2001; van Leijen, 2014).

2.4.3. InSAR

An interference pattern can be created by taking two SAR images from the same position in space but at separate points in time. In practice, two SAR images are taken and the measured phases are subtracted. The interference pattern, or fringes, is an indication of the phase difference between these two images. The distance between the two acquisition locations is called the baseline and should be small since, otherwise, geometric decorrelation will occur. The goal is to use the interference pattern to estimate the deformation in a particular area that happened within the period between the two acquisition moments. However, it is not possible to immediately use the obtained phase differences and compute the deformation. The measured signal has many contributing sources or phase contributors.

Interferometric phase contributors

The computed InSAR image contains phases composed of multiple phase contributors. The following equation shows the different phase contributors:

$$\phi^{ms} = \Phi^m - \Phi^s = -2\pi a + \phi_{\text{flat}} + \phi_{\text{topography}} + \phi_{\text{deformation}} + \phi_{\text{atmosphere}} + \phi_{\text{scattering}} + \phi_{\text{noise}}. \quad (2.23)$$

The equation shows the subtraction between the first image (Φ^m, master) and the second image (Φ^s, slave). Eq. (2.23) shows the seven phase contributors. The first contributor is a. a is the

integer number of cycles for the two-way distance. The next contriubutor is the Flat Earth Phase (ϕ_{flat}). The flat Earth phase describes the contribution due to a reference surface, e.g., an ellipsoid (**?**). The topographic phase ($\phi_{\text{topography}}$) describes the influence of the topography above the reference surface. The deformation phase ($\phi_{\text{deformation}}$) consists of the displacements of the surface and is the phase contributor of interest in this thesis. The atmospheric contributor ($\phi_{\text{atmosphere}}$) incorporates the delay the atmosphere introduces. The scattering contributor ($\phi_{\text{scattering}}$) and the noise contributor (ϕ_{noise}) are effectively the same, but the scattering contributor describes the changing backscatter characteristics of a pixel. The noise term consists of the processing noise introduced (van Leijen, 2014).

Deformation Phase

The deformation phase is the contributor of most interest, because it consists of the displacements on the surface. However, the phase measurement is the projection of the actual three-dimensional deformation vector onto the line of sight vector. The actual three-dimensional deformation vector is

$$d_{geo} = [d_e, d_n, d_u]^T. \tag{2.24}$$

The vector consists of the deformation in the East- (d_e), North- (d_n), and Up-direction (d_u). The relationship between the line of sight measurement and the three-dimensional vector is given by

$$d_{los} = p^T d_{geo} = [-\sin\theta \cos\alpha_h, \sin\theta \sin\alpha_h, \cos\theta] d_{geo}. \tag{2.25}$$

Eq. (2.25) makes use of the heading (α_h) of the satellite (the direction the satellite is flying w.r.t. north direction) and the incidence angle, θ (Chang et al., 2014; Hanssen, 2001). The incidence angle is inclined and therefore the satellite is sensitive to horizontal and vertical displacements. However, depending on the heading angle (α_h), the satellite is insensitive parallel to the heading angle and most sensitive perpendicular to the heading angle (Hanssen, 2001; van Leijen, 2014).

Furthermore, the phase contributor is still in cycles and displacements are preferably shown with a distance unit such as meters. The equation converting from cycles to meters, as described in van Leijen (2014), is given by

$$\phi_{deformation} = \frac{-4\pi}{\lambda} D_{los}. \tag{2.26}$$

D_{los} in Eq. (2.26) is the displacement in the line of sight. λ is the wavelength used by the radar, and $\phi_{deformation}$ is the displacement in cycles. The whole process of creating InSAR images is commonly automatized and described in the next section.

InSAR processing

In practice, the InSAR image creation is automatized, and a few extra steps are added to improve the quality of the data. These steps are visualized in the flowchart in Fig. 2.6. This thesis will not go further into the details of InSAR processing. For a more in-depth explanation of the various steps in the processing scheme, see Hanssen (2001); van Leijen (2014). With the InSAR processing scheme set-up, a stack of interferometric images can be created. These images contain temporal changes of scatterers. However, not every scatterer behaves consistently over time and tends to become incoherent or decorrelated. In order to solve this problem, an algorithm is designed to look for persistent scatterers in an interferometric data stack.

2.4.4. Persistent Scatterer Interferometry

Making multiple interferograms from the same position in space but at multiple moments (epochs) in time generates an interferometric data stack. The decorrelation of a scatterer in time is a problem when creating these InSAR time series. The Persistent Scatterer Interferometry (PSI) algorithm selects pixels in the time series that constantly behave over time. These pixels are stable scatterers. Two types of scatterers can be distinguished in an InSAR time series. The first type is the Point Scatterer (PS) which looks at one strong and stable reflection of an object over time. The second type of scatterer is the Distributed Scatterer (DS), where the distributed scattering mechanism is constructed by the sum of many independent small scatterers within a predefined resolution cell (Goel & Adam, 2013).

The PSI algorithm is constructed out of three steps. The first step in the Persistent Scatterer algorithm is the selection of Persistent Scatterer Candidates (PSC). This step is implemented to reduce

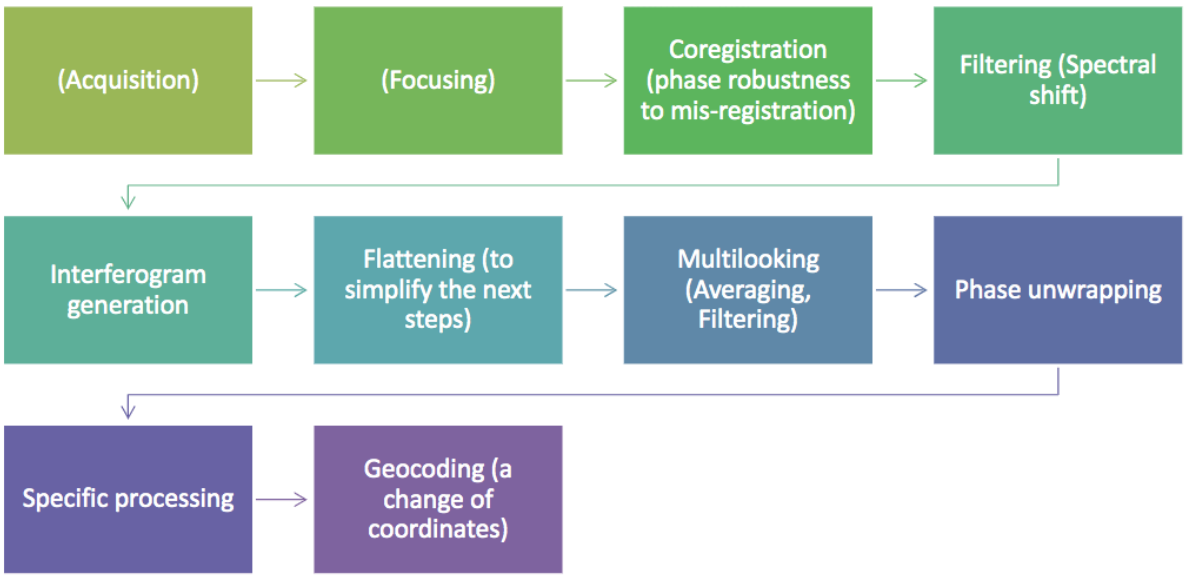


Figure 2.6: InSAR processing flowchart from López Dekker et al. (2018).

the amount of processed data and to apply a filter for the quality of the reflection points. The selection is based on the normalized amplitude dispersion D_a

$$D_a = \frac{\sigma_a}{m_a}. \tag{2.27}$$

Here, σ_a is the standard deviation of the amplitude, and m_a is the mean amplitude. The pixel selection threshold which is commonly used is $D_a < 0.25$ (Bruna, 2020), but it can be adjusted to the user's preference.

The second step is the 'First-order network estimation.' This step takes the most coherent scatterers and estimates displacement parameters and DEM errors. The obtained phase residuals enables the estimation of the atmospheric signal delay.

The third step is the 'Second-order network integration.' Here, the parameters at the selected point locations are obtained through a least-squares (LS) integration w.r.t. a reference point. From the second-order network, it is possible to derive an entire deformation velocity field.

2.5. Early Detection and Warning System

Impending sinkhole risk prediction and assessment requires identifying existing sinkholes as the first step and then making predictive statements where new impending sinkholes are likely to exist. The critical characteristics, which should be taken into account, are size, frequency, subsidence mechanisms, and rates (Gutiérrez et al., 2007). However, sinkholes are frequently masked by anthropogenic activities. Sinkholes commonly have a subtle surface expression, or the underground cavity is not close enough to the surface (Gutiérrez et al., 2007). The prediction of sinkhole occurrence can be subdivided into a temporal and a spatial prediction.

The temporal prediction has two different methods to predict future sinkholes (Gutiérrez et al., 2007). The first prediction method is based on an exact time in the future with a specific time interval (range). The second prediction method predicts the frequency or the probability of a sinkhole occurrence. Both methods heavily rely on the historical data that is available for a specific area. Areas with short or limited data tend to produce worse predictions. Also, when using historical data, a bias is introduced. The bias is produced by the time the surveys were undertaken, and there is a bias towards larger-sized sinkholes since they are much more often noticed and recorded (Gutiérrez et al., 2007).

The spatial prediction methods create susceptibility maps by looking at the locations of known sinkholes. Within the susceptibility maps, zones are defined with each their own level of importance. Zones with a higher level of susceptibility are close to or between the sinkhole occurrences (Gutiérrez et al., 2007). Another indication of increased sinkhole risk is identifying the overall direction of the newly

occurring sinkholes (Gutiérrez et al., 2007). These different levels of susceptibility and the dominant direction are indications of where to apply the monitoring and to what degree.

Jones & Blom (2014) suggests that for hazard monitoring, InSAR is an economically viable technique. The reason being that InSAR is the most feasible solution when monitoring a large area. InSAR produces wide imaging swaths and consistent repeat intervals. Therefore, an early warning system for impending sinkholes should be centered around the usage of InSAR data. With InSAR, a generic subsidence pattern can be discovered. By quantifying the surface strains, predictions can be made on the size and growth rate of the sinkhole.

Currently, at the Dead Sea, a new early warning system is in development. Baer et al. (2018) and Nof et al. (2019) describe an early warning system whereby the processing steps (i.e., satellite image acquisition to processed interferogram) are automatized. When the processing is finished, a specialist will inspect and compare the image to look for significant subsidence. The next step to further automatize this system is to apply machine learning to automatically detect newly forming sinkholes. Intrieri et al. (2015) used a ground-based SAR (GB-InSAR) device to monitor the area. Here, the paper predicted the location of an impending sinkhole before it could reach the surface. Significant subsidence was detected from the precursory subsidence. The sinkhole was predicted in advance with sufficient time to dispatch a warning and thus showed the feasibility of the early warning system.

Nof et al. (2019) states that the arid and sparsely populated Dead Sea is ideal when monitoring for sinkholes using interferograms. Less ideal areas are more inclined to use Persistent Scatterer Interferometry (Section 2.4.4). Malinowska et al. (2019) used PSI time series to detect and monitor ground movements potentially associated with the existence of an impending sinkhole. It was found that using PSI, it was still possible to identify zones where sinkhole would potentially occur, given adequate spatio-temporal sampling. The identification of these zones was achieved by looking at accelerated ground deformation. Accelerated ground deformation was detected within a radius of 100 meters of the impending sinkhole. Chang & Hanssen (2014) also discovered two acceleration events before the catastrophic sinkhole collapse in shopping mall 't Loon. The shopping mall is located in the Netherlands. The collapse occurred on December 3, 2011.

Using InSAR for early detection has its advantages but also has its limitations. Intrieri et al. (2015) identified two major limitations in using InSAR:

- The small size of a sinkhole with respect to the susceptible area to be monitored;
- The short time between the detection of precursors (if any) and the collapse.

These limitations are inline with the conditions mention in Section 2.3. The resolution determines the minimum detectable sinkhole size. The PSI technique looks for coherent scatterers. Coherent scatterers are obtained based on the surface cover type. Areas containing less ideal scattering surface covers provide less coherent scatterers. The result is that displacement estimates are not homogeneously distributed over the areas. A consequence is that the detection of a sinkhole is limited to the vicinity or the spacing between coherent scatterers. Impending sinkholes closer to a measurement point will show a larger displacement signal. More measurement points are important for determining the sinkhole size.

Another limitation is the acquisition time. Depending on the used satellite mission, the sinkhole must have a subsidence period longer than the acquisition time interval (Nof et al., 2019). However, with the advent of more frequent-repeat SAR missions, the possibility of detecting impending sinkholes will increase.

3

Methodology

In this chapter we discuss the detection methodology developed in this study. We begin by explaining the methodology developed for describing the surface expression. We can base the surface expression on the geometric model, which is Eulerian-based. A Lagrangian-based model is the Mogi Point Source Model and we will explain how we implemented this model. Adding a temporal aspect to the geometric model will create the kinematic model. With the kinematic model we can use displacement estimates to calculate unknowns. Section 3.2 will talk about the methodology for time series characteristics determination. We discuss the method to superimpose a sinkhole signal on top of a real dataset, and methods to create maps showing the minimum detectable size, minimum detectable deformation and detectability power. Section 3.3 describes the methodology for implementing the spatio-temporal strategy. The last section, Section 3.4, describes the implementation of the temporal strategy based on arcs.

3.1. Sinkhole Surface Expression

In this section, we look at the surface signature of an impending sinkhole, using models gathered from literature.

3.1.1. Scaling Factor

In Section 2.2.2, we introduced the geometric model based on influence functions found in the literature. The geometric model based on the inverse Gaussian curve is the most popular model for subsidence modeling (Ren & Li, 2008):

$$k_z = \frac{1}{R^2} \exp\left(-\pi \frac{r^2}{R^2}\right), \tag{3.1}$$

where R is the radius of influence, which governs the surface extent of the sinkhole. The literature stated that the subsidence beyond R can be neglected (Ren & Li, 2008). The influence function is based on the normal distribution. The normal distribution is given by

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right). \tag{3.2}$$

In Eq. (3.1) R^2 is similar to the σ^2 in Eq. (3.2). One σ around the mean is equivalent to 68% of the curve and therefore one might think that it is also the case for the R^2 in Eq. (3.1). Fig. 3.1 shows that the comparison between the Gaussian influence function and the normal distribution is not valid. Therefore we conclude that the assumption that beyond the distance R the subsidence is insignificant, is valid.

The radial distance r from the center towards an arbitrary location within the subsidence trough, is defined as:

$$r = \sqrt{(x-x_0)^2 + (y-y_0)^2}, \tag{3.3}$$

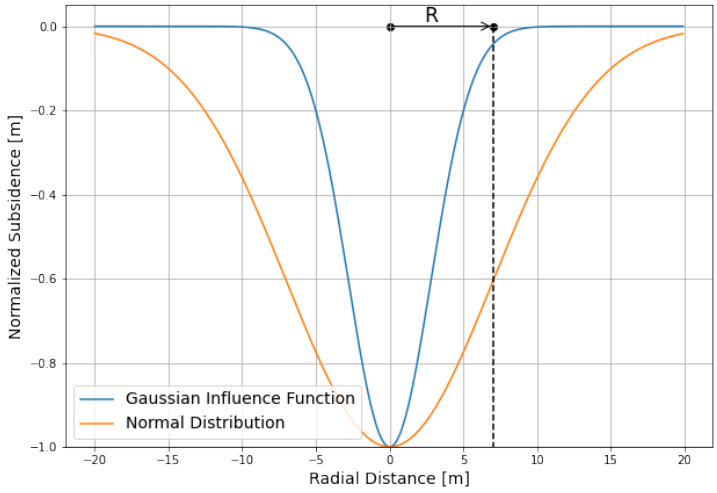


Figure 3.1: Comparison between the Gaussian influence function and the normal distribution. Both curves are normalized.

where x_0 and y_0 are the center coordinates of the depression of the impending sinkhole.

In Eq. (3.1) there is factor $\frac{1}{R^2}$ that is unnecessary and can be removed, i.e.,

$$k_z = \exp\left(-\pi \frac{r^2}{R^2}\right). \tag{3.4}$$

We can scale the geometric model to actual displacement values by multiplying the model of Eq. (3.4) with a scaling factor. We define the scaling factor as the maximum subsidence present at the center of the trough. Thus Eq. (3.1) with the scaling factor becomes:

$$v_z = S_{\max} \exp\left(-\pi \frac{r^2}{R^2}\right), \tag{3.5}$$

where $S_{\max}$ is the maximum subsidence and v_z becomes the vertical displacement.

In mining subsidence engineering (Kratzsch, 1983; Reddish & Whittaker, 1989) the maximum subsidence is related to the seam thickness, M, and the subsidence factor, a, with $S_{\max} = aM$. The subsidence factor is a dimensionless factor that is empirically found for the local overburden. Reddish & Whittaker (1989) described an empirical relationship used in the USSR as

$$S_{\max} = \frac{2Mw}{w + w_c}, \tag{3.6}$$

where $S_{\max}$ is the maximum subsidence, M is the seam height, w is the seam width and w_c is the critical seam width $w_c = 1.15h$, where h is the seam depth. The factor 1.15 follows from the limit angle, γ (see section 2.2.2), with $w_c = \frac{h}{\cos(\gamma)}$ (Reddish & Whittaker, 1989).

Hence, we use the scaling factor to tune the generic geometric model to specific cases. An impending sinkhole causes the ground to subside gradually and thus deepen the subsidence trough. The geometric model does not incorporate the temporal deepening of the trough. Therefore, we modify the scaling factor to fit our needs. We redefine the scaling factor as a subsidence velocity v and a time parameter t.

3.1.2. Kinematic Model

Modifying $S_{\max}$ to incorporate the subsidence velocity v and the time t gives:

$$S_{\max}(t) = v \cdot t. \tag{3.7}$$

Combining Eq. 3.7 and Eq. 3.4 gives:

$$v_z = v \cdot t \cdot \exp\left(-\pi \frac{r^2}{R^2}\right). \tag{3.8}$$

With Eq. 3.8 we can compute the subsidence (v_z).

Sinkholes, however, exhibit not only a linear behavior over time, but the subsidence can also accelerate in time.

$$v_z = a \cdot t^2 \cdot \exp\left(-\pi \frac{r^2}{R^2}\right) \tag{3.9}$$

whereby a is the acceleration and t is the time parameter. By making the geometric model time dependent, we convert the geometric model to a kinematic model.

3.1.3. Eulerian versus Lagrangian Displacement Comparison

We introduced two viewpoints in Section 2.2.2. The difference between the Eulerian and the Lagrangian viewpoint lies in the assumptions. The Eulerian viewpoint only concerns the elevation changes at a particular horizontal position, while the Lagrangian viewpoint tracks a point through space and time. To mimic both viewpoints and their conclusions, we compare two types of models. The first model is the geometric model, which is Eulerian. The Lagrangian physical model is the Mogi point source model (see Section 2.2.2) and modeled after a physical process.

The geometric model approximates the shape of the subsidence trough. The model has two parameters to adjust the shape: the radius of influence R and the maximum subsidence ($S_{\max}$). R influences the width of the trough, and $S_{\max}$ the vertical displacement. The disadvantage of using an Eulerian geometric model is that it does not compute any horizontal displacement. We know that both vertical and horizontal displacement is present during the precursory subsidence stage of a sinkhole. We can compute the horizontal displacement, but we have to base it on the vertical displacement and another assumptions, see Section 2.2.3.

Eq. (2.3) describes the volume-based Mogi source model. The important parameters are the Poisson's ratio, ν, the location of the source (x_0,y_0,z_0), and the volume change, ΔV. With these parameters, we compute the forces on each point in 3D space. We set the surface to $z = 0$, and thus the source will have the coordinates x_0, y_0, and $z_0 = -d$, where d is the depth of the source. This depth is computed using Eq. (2.4). In this case, d is the same as H. We compute H with R (defined for the geometric model) and θ, the angle of draw, which is set to $35°$ (Ren & Li, 2008).

Methodology for Comparing the Models

We described the geometric model and the physical model corresponding to the two Eulerian and Lagrangian viewpoint. Now, we explain how we compared both models.

We start by comparing the vertical displacement (subsidence) since the geometric model is based on subsidence calculation. We define a surface consisting of 300 points with the extent going from -20 meters to 20 meters. The center point is set to 0. We set the Poisson's ratio (ν) to 0.25 as suggested by Bekendam & Pottgens (1995) for Limburg. We apply for R an arbitrary value of 15 meters, and therefore fixing d. We compute d according to Eq. (2.4) since we know R and θ. The volume for the Mogi point source model is computed as follows

$$V_{\text{volume}} = \frac{4}{3}\pi r^3. \tag{3.10}$$

We compute ΔV by Eq. (3.10). The parameter r in this equation is the radius of the sphere. We set r to an arbitrary value of 0.5 meters creating a volume of slightly more than 0.5m^3.

We then take the maximum absolute value of the displacement vector d_u and assume it to be the maximum subsidence for the kinematic model. In this way, we can scale the geometric model to the physical model.

Next, we compare the horizontal displacements of the geometric model and physical model. We compare the horizontal displacement of the physical and the kinematic model. We obtain the horizontal displacement of the kinematic model with the focal point method (see Section 2.2.3). We already know the depth of the cavity, since we need that information for the physical model.

The last comparison we make between the two viewpoints is how they behave temporally. We make the comparison by consistently subtracting the same amount of volume, i.e., $-\Delta V$, for the physical model and use its maximum subsidence to scale the kinematic model. We update the x, y, and z coordinates at each new time step.

3.1.4. Estimating kinematic model parameters using BLUE

We commonly perform computing unknown parameters, or estimates, with geodetic measurements by Best Linear Unbiased Estimation, or BLUE. With this technique, we define a functional model, $E\{\underline{y}\} = Ax$, together with a stochastic model, $D\{\underline{y}\} = Q_y$. The *design matrix* A describes the functional relationship between the observations, $\underline{y}$, or measurements, and the estimates x. The stochastic properties are parameterized by the variance-covariance matrix, Q_y, of the measurements.

We solve the functional model, finding the Best Linear Unbiased Estimator, using:

$$\underline{\hat{x}} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} \underline{y}. \tag{3.11}$$

With an estimation for $\underline{\hat{x}}$, the adjusted observations, $\underline{\hat{y}}$, and the residuals, $\underline{\hat{e}}$, are:

$$\underline{\hat{y}} = A\underline{\hat{x}}, \tag{3.12}$$

$$\underline{\hat{e}} = \underline{y} - \underline{\hat{y}}. \tag{3.13}$$

The benefit of using weighted linear least-squares is that it incorporates the stochastic properties of the observations. We calculate the stochastic properties of the estimators using:

$$Q_{\hat{x}} = (A^T Q_y^{-1} A)^{-1}, \tag{3.14}$$

and the stochastic properties for the adjusted observations and residues are:

$$Q_{\hat{y}} = A Q_{\hat{x}} A^T, \text{ and} \tag{3.15}$$

$$Q_{\hat{e}} = Q_y - Q_{\hat{y}}. \tag{3.16}$$

In this study we use weighted linear least-squares for the spatio-temporal strategy to compute specific unknown parameters for the kinematic model. The kinematic model was expressed by Eqs. (3.8) and (3.3). We can dissect the maximum subsidence parameter of the kinematic model into a velocity v (or acceleration a) and a time parameter t, see Eqs. (3.7) and (3.8). We predominantly use the linear relationship throughout this study. In order to apply weighted linear least-squares to the kinematic model, we have to define and assume a few parameters first. Starting with the radial distance, see Eq. (3.3), the x and y parameters are the coordinates of the measurement points. The x_0 and y_0 coordinates are the center of the phenomenon, i.e., the depression preceding a sinkhole. The location of the sinkhole is not known a priori, and thus we do not know x_0 and y_0. However, we can implement a grid search approach. We define a grid over our area of interest. Then, on each grid point, or posting, we assume that it is the center of the impending sinkhole, defining x_0 and y_0. We define the time parameter t as the vector of all the used epochs in days relative to the first epoch, e.g., $t = [0, 6, 12, 18]$.

The next parameter is the radius of influence, R. The size of the impending sinkhole is not known a priori. There are two approaches to handle the parameter R. The first approach is to estimate both the radius of influence and the subsidence velocity. This means that the kinematic model becomes a non-linear function when solving for R and v. The kinematic model has to be linearized and iteratively solved.

The second approach is to simply assume a value for R. By assuming a value for R, the kinematic model becomes linearly solvable.

We will first discuss how we linearize the non-linear kinematic model to solve for R and v. Next, we will discuss how assuming R would make the kinematic model linear and thus linearly solvable for v.

Estimating the radius of influence and the subsidence

This section will explain how we linearize the Gaussian kinematic model and estimate the two unknowns.

The vertical displacement

$$E\{z(r)\} = Se^{-\pi \frac{r^2}{R^2}} \tag{3.17}$$

is a function of the radial distance, r, from the center of the bowl, see Eq. (3.3). The scaling factor S is a function of time t. As time t increases, the Gaussian bowl will deepen. Here, we assume a linear relation

$$S(t) = vt, \tag{3.18}$$

and thus Eq. (3.17) translates to

$$E\{\underbrace{z(t,r)}_{y}\} = \underbrace{vte^{-\pi\frac{r^2}{R^2}}}_{A(x)=A(v,R)}. \tag{3.19}$$

There are two unknowns in the forward model, i.e., v and R. However, Eq. (3.28) needs to be linearized with respect to the two variables that need to be determined.

$$E\{\underbrace{\Delta z(t,r)}_{\Delta y}\} = \underbrace{\begin{bmatrix} a_1 & a_2 \end{bmatrix}}_{\partial_x A(x^0)} \underbrace{\begin{bmatrix} \Delta v \\ \Delta R \end{bmatrix}}_{\Delta x}, \tag{3.20}$$

where

$$a_1 = \frac{\delta z}{\delta v} = te^{\frac{-\pi r^2}{R^{0^2}}}, \text{ and} \tag{3.21}$$

$$a_2 = \frac{\delta z}{\delta R} = \frac{2\pi r^2}{R^{0^3}} vte^{-\pi\frac{r^2}{R^{0^2}}} = \frac{2\pi r^2}{R^{0^3}} z^0(r), \tag{3.22}$$

where z^0 follows from fitting the initial values for R^0 and v^0 in Eq. (3.20).

If we have N observation points at T epochs, we get

$$E\{\Delta z(t,r)\} = E\{\begin{bmatrix} \Delta z(t_1,r) \\ \Delta z(t_2,r) \\ \vdots \\ \Delta z(T,r) \end{bmatrix}\} = \underbrace{\begin{bmatrix} a_1(t_1,r) & a_2(t_1,r) \\ a_1(t_2,r) & a_2(t_2,r) \\ \vdots & \vdots \\ a_1(t_T,r) & a_2(t_T,r) \end{bmatrix}}_{A_\circ} \begin{bmatrix} \Delta v \\ \Delta R \end{bmatrix} \tag{3.23}$$

where

$$\Delta z(t_i,r) = \begin{bmatrix} \Delta z(t_i,r_1) \\ \Delta z(t_i,r_2) \\ \vdots \\ \Delta z(t_i,r_N) \end{bmatrix}, \; a_1(t_i,r) = \begin{bmatrix} a_1(t_i,r_1) \\ a_1(t_i,r_2) \\ \vdots \\ a_1(t_i,r_N) \end{bmatrix}, \text{ and } a_2(t_i,r) = \begin{bmatrix} a_2(t_i,r_1) \\ a_2(t_i,r_2) \\ \vdots \\ a_2(t_i,r_N) \end{bmatrix}. \tag{3.24}$$

Given the initial values R^0 and v^0, we find

$$\Delta z(t_i,r) = z(t_i,r) - A(x^0), \tag{3.25}$$

where we find $A(x^0)$ from Eq. (3.20) by filling in the initial values. We then estimate the unknown parameter via

$$\begin{bmatrix} \hat{v} \\ \hat{R} \end{bmatrix} = \begin{bmatrix} v^0 \\ R^0 \end{bmatrix} + (A_\circ^T Q_y^{-1} A_\circ)^{-1} A_\circ^T Q_y^{-1} \Delta z(t,r) \text{ and} \tag{3.26}$$

$$Q_{\hat{x}} = (A_\circ^T Q_y^{-1} A_\circ)^{-1}. \tag{3.27}$$

We estimate the unknown parameters iteratively using the Gauss-Newton iteration method. With this method, we create a loop that converges to a particular answer.

Estimating the subsidence velocity

Estimating only the subsidence velocity is a linear problem and thus it doesn't use the iterative scheme we showed in the previous section. Another benefit is that we can use the same linear least-squares approach for all geometric models. It is quicker and more reliable, but we have to make an extra assumption assuming a value for R.

The vertical displacement in the new case is still Eq. (3.28) with r explained by Eq. (3.3). The scaling factor S is again rewritten as a linear relationship between subsidence velocity v and time t and shown by Eq. (3.18). The final equation translates to

$$E\{\underbrace{z(t,r)}_{y}\} = \underbrace{vte^{-\pi\frac{r^2}{R^2}}}_{A(x)=A(v)}. \tag{3.28}$$

This forward model only has one unknown and is linear. Therefore, we use the Best Linear Unbiased Estimator (BLUE). The equation in matrix notation will look like

$$E\{\underbrace{z(t,r)}_{y}\} = \underbrace{[te^{-\pi\frac{r^2}{R^2}}]}_{A}\underbrace{[v]}_{x} = [a_1(t,r)][v] \tag{3.29}$$

If we have N points at T epochs, we get

$$E\{z(t,r)\} = E\{\begin{bmatrix} z(t_1,r) \\ z(t_2,r) \\ \vdots \\ z(t_T,r) \end{bmatrix}\} = \underbrace{\begin{bmatrix} a_1(t_1,r) \\ a_1(t_2,r) \\ \vdots \\ a_1(t_T,r) \end{bmatrix}}_{A}[v], \tag{3.30}$$

where

$$z(t_i,r) = \begin{bmatrix} z(t_i,r_1) \\ z(t_i,r_2) \\ \vdots \\ z(t_i,r_N) \end{bmatrix}, \text{ and } a_1(t_i,r) = \begin{bmatrix} a_1(t_i,r_1) \\ a_1(t_i,r_2) \\ \vdots \\ a_1(t_i,r_N) \end{bmatrix}. \tag{3.31}$$

We then estimate the unknown parameter via

$$\hat{v} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} z(t,r) \tag{3.32}$$

and the stochastic properties of the estimate are obtained by

$$Q_{\hat{v}} = (A^T Q_y^{-1} A)^{-1}. \tag{3.33}$$

Stochastic Properties of measurements

We use the variance-covariance matrix Q_y in both the linearized case and the linear case. However, these measurement variances and covariances, i.e., the precision, are not known, and only rough estimates are available, which may not be reliable. Therefore, we write the stochastic matrix as a weight matrix with equal weights on the diagonal. The relationship is as follows:

$$W = c\, I \tag{3.34}$$

In most cases, we assume the measurements to be uncorrelated spatially and temporally since every measurement is independent of the other. Strictly speaking, a correlation can be present since large effects, such as the atmosphere or large-scale surface motion, may influence every measurement.

3.1.5. Defining error metrics to evaluate the performance of multiple models

We analyze the displacement curves by fitting the kinematic model through them. We assess the performance of the models by two error statistics and a signal-to-noise comparison. Here, we discuss three metrics. The first metric is the *fit* and is denote as f:

$$f = 100 \cdot \left(1 - \sqrt{\frac{\hat{e}^T \hat{e}}{y^T y}}\right). \tag{3.35}$$

With Eq. (3.35) the percentage of goodness of fit of the model to the measurements is computed. The closer the fit is 100%, the better the fit is. The next metric is the Root Mean Square Error (RMSE). The RMSE is given by

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}(\underline{y} - \hat{y})^T(\underline{y} - \hat{y})}. \tag{3.36}$$

Eq. (3.36) is also known as the root mean square deviation (RMSD). The RMSE statistic is a measure of the performance of a model. We apply a term-by-term comparison of the difference between the observations and the adjusted observations—the smaller the value, the better the model's performance. However, a large drawback of the RMSE test is that a few larger errors (outliers) can significantly influence the RMSE outcome (Kambezidis, 2012).

Since we do not know the exact center of the impending sinkhole, we use a range of impending sinkhole centers. We also use the RMSE to identify the best possible impending sinkhole center location (x_0). An approximation for the trough center is visually inferrable from the figures of the cases. However, we prefer a more precise estimation based on a quantifiable metric. Therefore, we evaluate multiple sinkhole centers. We then select the center producing the lowest RMSE for the comparison. We chose the RMSE since the papers used (Kim et al., 2016, 2019; Nof et al., 2019) also applied it to assess the performance of their models.

The last metric is the signal-to-noise (SNR) comparison. Here, we make a comparison between the adjusted observation and the residuals.

$$\mathrm{SNR} = 10 \cdot \log_{10}\left(\frac{\hat{y}}{\hat{e}}\right) \tag{3.37}$$

We can see the SNR in Eq. (3.37) in decibels (dB). We compute the SNR for each position. Therefore, we can observe in what locations the expected signal is the strongest and how the signal weakens when going toward the edges of the trough. Also, we implement a threshold. We first take an arbitrary value for the threshold. This value was set to 10 dB, meaning that the signal is approximately ten times higher than the noise. Then, we compute the number of points above the arbitrary threshold, creating a new metric that computes what percentage is above the threshold, and thus *measurable*.

3.1.6. Testing the Kinematic Model on Real-world cases

The kinematic model is mostly theory-based, and explanations of this theory can be found in Section 3.1, Kratzsch (1983) and Reddish & Whittaker (1989).

The theory does not explicitly describe the use of the kinematic model on impending sinkholes. Therefore, we want to validate whether the model represents the impending sinkhole surface expression adequately. Hence, we apply the kinematic model to a few cases found in published papers. The displacement time-series where found in the papers Kim et al. (2016), Kim et al. (2019) and Nof et al. (2019), and subsequently digitized.

Many natural phenomena follow the Pareto distribution. The Pareto distribution is also known as the Pareto principle or the 80/20 rule, and we can describe it with a power-law distribution. The Pareto principle states that 80% of the consequences can be attributed to 20% of the causes. In the case of the kinematic model, the assumption is that the model would describe roughly 80% of the curve (Dunford et al., 2014). We test this assumption on displacement curves digitized from papers. We digitize the displacement time series from these papers and subsequently use the displacement time series to estimate the parameters for the kinematic model. We aim to test whether the kinematic model would adequately describe the displacement pattern regarding the Pareto distribution.

Error Metrics

We apply the different models to the digitized displacement time series. We then compute the overall score by taking the average of the *RMSE* and *fit* over every epoch (Section 3.1.5). However, the average fit parameter does not include the first epoch. Due to the models, the residuals correspond one-to-one with the data and therefore generate a zero fit. Thus, we exclude the first epoch in the computation of the average fit of a model.

Implementation Kinematic Model

We digitized a temporal cross-section of a sinkhole. We need to find the center (x_0) and the corresponding subsidence velocity and the radius of influence. We estimate these parameters using the non-linear least-squares method described in Section 3.1.4. We test multiple locations on the cross-section line and we select the estimate providing the lowest RMSE.

Implementation Mogi Source Model

To fit the Mogi model, we had to adjust the model (Eq. (2.3)) to incorporate a time parameter. As mentioned in an earlier section, the Mogi model is Lagrangian-based and thus follows specific points from its origin. We digitized the data from papers. The papers did not show specific points (or the points were not always individually distinguishable) and therefore we lose point location information. We therefore assumed that for each progressing epoch, the volume change (ΔV) increases and that at each new epoch, the point start from their original position.

Adjusting the Mogi model gives:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = t \cdot \Delta V_i \frac{1-\nu}{\pi} \begin{bmatrix} \frac{x}{R^3} \\ \frac{y}{R^3} \\ \frac{z}{R^3} \end{bmatrix}. \tag{3.38}$$

The cross-section is in 2-dimensions and this makes:

$$\begin{bmatrix} u \\ w \end{bmatrix} = t \cdot \Delta V_i \frac{1-\nu}{\pi} \begin{bmatrix} \frac{x}{R^3} \\ \frac{z}{R^3} \end{bmatrix}. \tag{3.39}$$

On Eq. (3.39) we can apply the linear least-squares to estimate ΔV_i.

We apply the same methodology as for the kinematic model. We define a range of x_0 values along the cross-section and compute the estimate. The estimate with the lowest RMSE will be selected for comparison.

3.2. Spatial and Temporal Time Series Characteristics

The second sub-question tries to look at the characteristics of the dataset. Since our data does not contain a sinkhole, we first discuss how we implement an artificial sinkhole into the dataset using the kinematic model (Section 3.2.1). The following subsections will explain methods for computing the minimal detectable sinkhole size (Section 3.2.2) and the method to compute the minimal detectable deformation given a detectability power and significance level (Section 3.2.4).

3.2.1. Artificial Sinkhole Implementation

The sinkhole collapse event is a high-impact, low-probability event. This means that obtaining data containing an active impending sinkhole, is not always possible. In this section we will explain how an impending sinkhole signal is added to the data to simulate an impending sinkhole event.

The kinematic model is used to simulate the impending sinkhole signal (see Section 3.1.2). Then, to implement the model, we need to define the location, radius of influence, the time duration of the event and the maximum displacement. A well documented sinkhole case is the sinkhole at shopping mall 't Loon. A near-collapse sinkhole appeared at the end of 2011. The final near-collapse sinkhole had a diameter of ~8 meters. One well documented subsidence event shows that during the summer (~3 month period) displacements up to 2.1 centimeters was measured (Chang & Hanssen, 2014). This provides us with potential sinkhole characteristics based on real world measurements. However, for testing of the two strategies, a more visible sinkhole was opted. The implemented sinkhole has the following characteristics: The location of the sinkhole was such that at least one point laid close to the center and at least a few points are within the radius of influence of the impending sinkhole.

The defined kinematic model is superimposed on a subset of a real dataset. The research was conducted with the area of Limburg in mind. Here, we utilized a dataset processed by Skygeo. A subset of the dataset was created and used. The reason for using a real dataset is to keep the spatial distribution and the displacement time series as authentic as possible. Table. 3.1 shows the properties of the implemented impending sinkhole signal.

Table 3.1: Implemented Sinkhole characteristics. Based on these defined parameters, an artificial sinkhole is superimposed on the data.

Duration	6 months
Maximum Displacement	5 cm
Radius of Influence	35 m
Start Month	1 June 2018
End Month	31 December 2018

3.2.2. Minimum Detectable Size

The *minimum detectable size*, or MDS, shows the minimum size of an impending sinkhole that can be detected at a particular position. It uses the kinematic model to make this assumption. The MDS computes the smallest possible kinematic model possible to get an estimate. It shows how the spatial distribution of a particular dataset influences the detection possibilities. The MDS will therefore be applied to the real subset.

The steps for the method of the MDS are as follows:

1. Defining a grid on which each posting the kinematic model will be evaluated. A grid is created by simply taking the maximum and minimum values for the x- and y-coordinates in the dataset. The spacing (or posting) of the grid points is defined beforehand. The smaller the grid, the longer the map creation will take.

2. On each posting, a subset using the current radius of influence is created using the K-D tree data structure (see Section 3.2.2).

3. Using the available data points and the location of the center of the impending sinkhole (posting coordinates), the design matrix for the estimation of the subsidence velocity is created (see Section 3.1.4).

4. The condition number of the design matrix is computed see if the estimate is stable (see Section 3.2.2).

5. Step 2-4 are repeated for the range of R-values defined a priori.

Subset creation based on K-D tree data structure

K-D tree is short for K-dimensional tree. It is a binary tree that represents geometric data hierarchically. Here, the K-D tree is used to create subset rapidly. It was first described by Friedman et al. (1977). The data structure is optimal for range searching and nearest neighbor searching. The algorithm subdivides space into rectangular cells. This subdivision will continue recursively until only one point is present within the cell (Maneewongvatana & Mount, 1999). (Maneewongvatana & Mount, 1999) The time complexity for generating the K-D tree is $O(DN\log(N))$ and is a significant improvement over brute-force approach with large N. Fig. 3.2 show the visual cell subdivision of a two dimensional dataset.

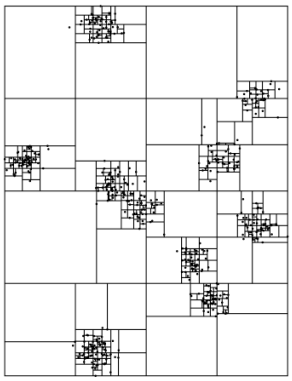


Figure 3.2: Visual representation of how the cell are defined when implementing the K-D tree data structure. The figure is copied from Maneewongvatana & Mount (1999) and shows the *Sliding-midpoint split*.

The splitting of the cells is governed by the *Sliding-midpoint split* and is also implemented in the Python module *scipy.spatial.cKDTree* (Virtanen et al., 2020).

Condition Number

The estimate computed using BLUE (see Section 3.1.4) might be highly sensitive to perturbations in the input data. A problem is said to be insensitive, or well-conditioned if a given relative change in the input data causes a reasonably proportionate relative change in the estimate. An sensitive problem, or ill-conditioned problem, is if a relative change in the estimate is much larger than that of the input data (Heath, 2018). The condition number of a problem is the ratio of the relative change in the estimate to the relative change in the input.

The definition for the condition number is given by

$$\text{cond}(A) = ||A|| \cdot ||A^{-1}||. \tag{3.40}$$

By convention, $\text{cond}(A) = \inf$ when A is a singular matrix. Eq. (3.40) makes use of the inverse of matrix A. A nonsquare matrix A, which happens a lot in linear least squares, does not have an inverse in the conventional sense. The solution is to define a pseudoinverse, denoted as A^+, and behaves like an inverse in many respects.

For linear least squares problem, the condition number is defined for a rectangular matrix.

$$A^+ = (A^T A)^{-1} A^T. \tag{3.41}$$

Trivially, we see that $A^+ A = I$. Then the new equation for the condition number becomes

$$\text{cond}(A) = ||A||_2 \cdot ||A^+||_2. \tag{3.42}$$

By convention, $\text{cond}(A) = \inf$ if $\text{rank}(A) < n$ whereby A is an $m \times n$ matrix with $\text{rank}(A) = n$. Just as the condition number of a square matrix measures closeness to singularity, the condition number of a rectangular matrix measures closeness to rank deficiency.

Evaluating the design matrix

Eq. (3.8) also uses a time vector. The time vector consists of epochs in days relative to the first epoch (e.g., $t = [0, 6, 12, 18]$). The number of epochs and the value of each epoch can be taken arbitrarily.

With the ingredients grid points, kinematic model and condition number, a quick map showing the minimal detectable size can be created. Such a map rapidly shows whether an early warning system would be feasible for the sinkhole hazard and typical sinkhole sizes prone to the area, which sensor is needed and what locations can be well monitored.

3.2.3. Quick and Simple Noise Determination based on displacement time series

This method is based on the displacement time series of either (coherent scattering) points or arcs (see Section 3.4). It uses the fact that there appears to be an overall displacement trend and the time series is *scattered* around that trend. We apply this method so we can get a rough estimate on the precision of the point and potentially predict displacement values to a certain degree.

An example of the displacement time series is shown in Fig. 3.3. Here, the line of sight displacement is shown on the y-axis and the epochs are scaled relative to the first epoch of the dataset. This means that the x-axis is in days since the first epoch. The reason for using relative days is so that a prediction of the trend can be made. The trend is computed using the linear least-squares. Here, a functional relationship between the observations and the unknowns must be defined. The unknowns are the velocity (a) and the bias (b). With n observations, the observation vector is

$$\underline{y} = \begin{bmatrix} d_{t_0} \\ \vdots \\ d_{t_n} \end{bmatrix}, \tag{3.43}$$

whereby the d is the LOS displacement at a particular epoch (t_i). The functional relationship is defined in the design matrix (A-matrix). Since we have two unknowns, the design matrix consists of two

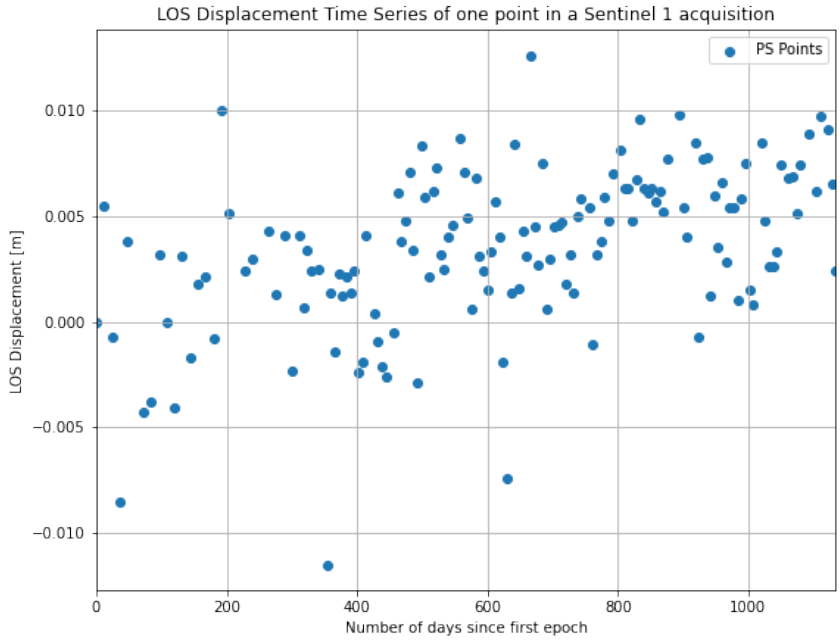


Figure 3.3: LOS Displacement time series of one single point scatterer.

columns and is written as

$$A = \begin{bmatrix} t_0 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix}. \tag{3.44}$$

Using BLUE (see Section 3.1.4) the velocity (a) and the bias (b) are computed.

In the computation of the BLUE, the weight matrix W is set to the inverse of the stochastic matrix of the measurements. However, the stochastic matrix in not known. Therefore, the weight for each measurement is set to one and results in

$$W = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} = I. \tag{3.45}$$

After obtaining the unknowns velocity (a) and the bias (b), the adjusted observation $\underline{y}$ can be computed

$$\underline{\hat{y}} = A\underline{\hat{x}}, \tag{3.46}$$

which is needed to compute the residuals. The residuals on their part can be computed using

$$\underline{\hat{e}} = \underline{y} - \underline{\hat{y}}. \tag{3.47}$$

Next, the assumption is made that the residuals follow a Gaussian distribution. Fig. 3.4 is the distribution of the residuals and shows the residuals plotted as a histogram with its probability density curve. The mean is assumed to be zero. The variance is then computed and used as the precision of the displacement time series of that particular point or arc.

3.2.4. Minimal detectable deformation and detectability power

Other rapidly created maps are the Minimal Detectable Deformation and the Detectability Power. The Minimum Detectable Deformation (MDD) provides a metric of how much deformation must be present in order to confidently attribute the deformation to a particular deformation signal w.r.t. a given Detectability Power (DP). In this section, the methodology of calculating these metrics will be discussed.

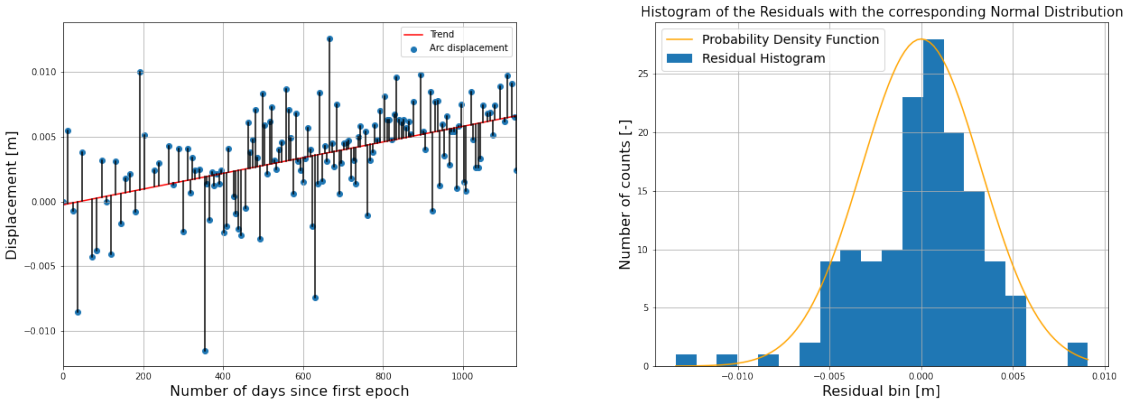


Figure 3.4: Left image shows a displacement time series with a linear trend in red and the residuals in black. The right image is the histogram of the residuals with the assumed Gaussian behavior plotted.

First, an observation time series is characterized by a linear model. The model assumes a standard deviation to each observation which acts as a proxy for the noise a particular observation exhibits. Then, the standard deviation is used for hypothesis testing which calculates the MDD and DP values.

The Minimum Detectable Deformation and the Detectability Power are calculated by testing hypothesis. For each new observation $\underline{y}_i$ we can apply the test. Two hypothesis are formulated, the first hypothesis is the null hypothesis (H_0). Here, the null hypothesis states that no significant deformation is present. The alternative hypothesis (H_a) states that a particular deformation is present. The hypotheses can be mathematically described as

$$H_0 : E\{y\} = x_0, \tag{3.48}$$

and

$$H_a : E\{y\} = x_a > x_0. \tag{3.49}$$

Then, the critical region must be specified. The critical region (K) is the region where H_0 will be rejected and H_a will be accepted. The assumption we take here is that each observation ($\underline{y}$) follows a Gaussian distribution. Since H_a states that the new expected value ($E\{y\}$) must be larger than in H_0, the critical region will only fall to the the right side of the distribution and becomes a one-tailed test. Fig. 3.5

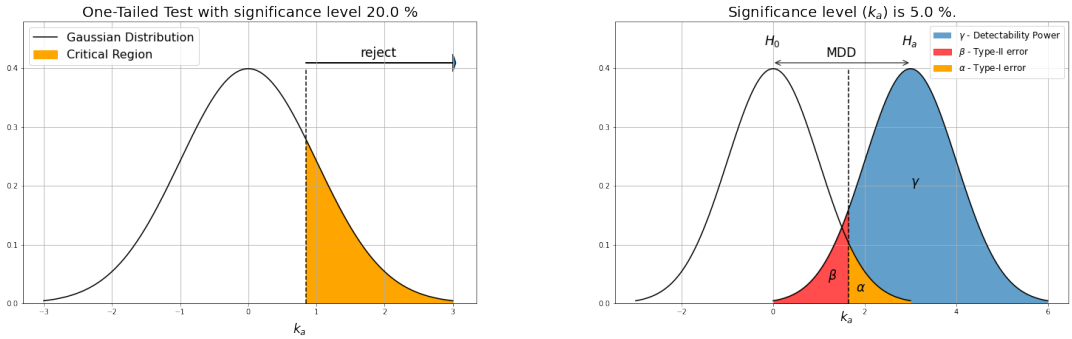


Figure 3.5: The left image is the Gaussian distribution of an observation $\underline{y}$. The orange area shows the critical region (K) where the null hypothesis will be rejected. The right image is a figure illustrating the hypothesis test.

shows how a particular critical region looks like and when observations will be rejected in favor of the alternative hypothesis. Fig. 3.5 shows both the null hypothesis and the alternative hypothesis. The critical region (orange region in Fig. 3.5) is governed by the significance level (α). The significance level manages the type-I error. The Type I error states that the null hypothesis is falsely rejected (i.e.,

false negative). The type II error is the error whereby the null hypothesis is falsely accepted (i.e.,false positive) and the region of this error is denoted with β (red region in Fig. 3.5). The Detectability Power (γ) is calculated by $\gamma = 1 - \beta$ and shown in Fig. 3.5 as the blue region. The Minimal Detectable Deformation is the distance between the expected value of the null hypothesis and the alternative hypothesis (Teunissen, 2000). Be aware that the claim whether a particular hypothesis test has been proven true can not be made (Teunissen, 2000).

3.3. Spatio-temporal Strategy using kinematic model

In this section, we take a closer look at the data. The most straightforward approach to detect impending sinkholes is looking for the sinkhole shape in the data. We approximate the sinkhole shape using the kinematic model (Section 3.1.2).

3.3.1. Grid-wise implementation

When we want to look for a sinkhole shape based on the kinematic model, we need to fit its function to the data. Eq. (3.8) describes the kinematic model whereby r consists of Eq. (3.3). The unknowns in the equations are x_0, y_0, R, and v. Looking at the equation, we see that this is not solvable in a linear fashion. To solve for the origin location of the sinkhole, we can implement a grid-wise search. This grid-wise search means defining a grid with a specific spacing or posting on top of the dataset (or subset). We then will systematically evaluate the equation on every grid position therefore predetermining the center coordinates x_0 and y_0.

3.3.2. Linear Least-Squares

However, R and v are still unknown, and the equation is still not solvable in a linear fashion. There are two ways we can solve this problem. The first option is to linearize the function and apply the Newton-Raphson method (Section 3.1.4). Second, we also evaluate a series of values for R at each posting. With the second approach, we can linearly solve the equation. In this methodology, we prefer the second approach. The reason is that the non-linear version might take considerably longer (due to the number of iterations) and might also produce an estimate that is not physically interpretable. For example, a subsidence velocity of -1 meter per day is not realistic.

3.3.3. Filter Results

The current approach defines a grid on which we apply the kinematic model on each posting. We have predefined R, and we can compute r by using the locations of the observations and the posting coordinates (x_0, y_0). The time vector t is the number of epochs used for the model and is in relative days. For example, we assume the impending sinkhole started 15 epochs ago. Therefore, the first epoch will get 0 (days), and each consecutive epoch will increase the number of days since the previous epoch. v_z is the displacement value corresponding with the observation location and epoch. We then solve for v, the subsidence rate or velocity.

The estimate we obtain for v ($\underline{\hat{x}}$) can still contain various values. We are only interested in a subsidence velocity within a specific range. Therefore, we filter the estimated rates such that only we accept rates within a prespecified range. Other values will receive a *NaN* value. *NaN* stands for Not a Number. We apply the filter to both the estimate ($\underline{\hat{x}}$) and the variance of the estimate ($Q_{\hat{x}}$).

3.4. Temporal Strategy using arcs

This section will explain the methodology behind the temporal strategy based on arcs. Arcs are the connections between points. The number of connections that can be made is

$$E = \frac{N(N-1)}{2}, \tag{3.50}$$

connections, whereby N is the number of points within the subset. This section will first explain how the number of arcs is defined and how we efficiently compute the connection pairs. We then characterize each arc by subtracting one endpoint from the other endpoint, creating a new displacement time series per arc. The next step is characterizing each arc with a (linear) model. A simple linear least-squares are applied to the time series to obtain a model predicting a trend. The model can be updated using different *update strategies*. Several *update strategies* will be discussed in this section. We then need a way to compute the *noisiness* of the time series. We compute the noisiness by first computing the residuals. We compute the residuals by taking the trend (provided by the model) or adjusted observations, $\underline{\hat{y}}$, and subtracting the observations, $\underline{y}$. The residuals are assumed to be Gaussian distributed and plotted in a histogram. We then use the standard deviation of this distribution as a proxy for the noisiness of the arc displacement time series. Finally, with the standard deviation and the trend, we can compute confidence intervals. We then test the newly acquired epoch whether it falls within the predefined confidence interval. We flag the arcs that failed the test. These steps are called the *flagging procedure*. The final step in the *flagging procedure* is to track the frequency on how often an arc is flagging in one row. We achieve this by setting flagged arcs to 1 plus the flagged arc value of the previous epoch. Therefore, the arc will obtain a number called *flagged value*.

After the flagging procedure, we can plot the flagged arcs and color the arcs according to the *flagged value*. However, the interesting information is to find the points with an increased number of flagged arcs towards them. The hypothesis is that an increase in flagged arcs will lead to the detection of anomalous behaving points. Therefore, we dismantle the arcs into points again corresponding to the input dataset. Each point will obtain the sum of the flagged value of each arc going towards a particular point. The *total flagged point value* will act as the measure of how anomalous a point is behaving. We apply a threshold to filter out these anomalous points, whereby we state that every point exceeding the threshold value behaves anomalously.

We explained, in short, the methodology of the temporal strategy. Below will be a more extensive explanation of this particular strategy.

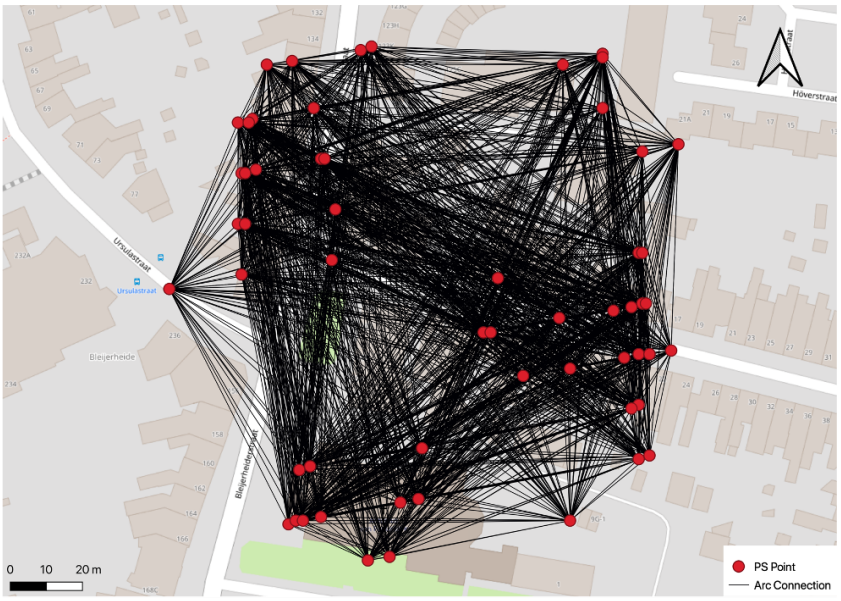


Figure 3.6: Figure showing all the possible arc connections within a subset. The figure shows that more observations can cause the figure to get computational expensive and messy.

3.4.1. Arc Creation

The main motivation of the usage of arcs is to eliminate most error sources commonly found in every observation and to optimally exploit the information content of each point. By subtracting two points from each other, we eliminate the correlated errors affecting those two points.

We compute the arcs by creating tuples (p_i, p_j) of combinations of points (two points denoted here as p_i and p_j) within the total number of points (N). One way to create all the tuples for all the arcs is by simply going from point to point and create all the tuples connecting that point. However, this introduces many duplicate tuples and is therefore very inefficient. Here, we implemented a tuple creating scheme based on a lower triangular matrix to identify all possible combinations. For example,

Table 3.2: Matrix for defining the arc combinations used for the temporal anomaly detection strategy.

	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0
3	1	1	0	0	0	0	0	0	0	0
4	1	1	1	0	0	0	0	0	0	0
5	1	1	1	1	0	0	0	0	0	0
6	1	1	1	1	1	0	0	0	0	0
7	1	1	1	1	1	1	0	0	0	0
8	1	1	1	1	1	1	1	0	0	0
9	1	1	1	1	1	1	1	1	0	0
10	1	1	1	1	1	1	1	1	1	0

a subset with $N = 10$ points will create the Table 3.2. It shows the lower triangular for a subset of ten points. We automated the creation of this tuple list with a Python function (See Appendix A).

3.4.2. Processing Arcs

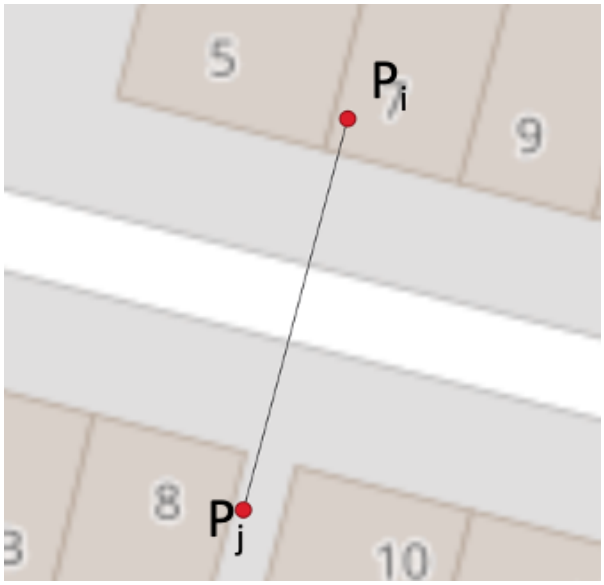


Figure 3.7: Figure showing two points with the arc they are making. The points are denoted as P_i and P_j.

We now know how to compute all arcs in a subset efficiently. Here, we explain what the procedure is to compute the new displacement time series of each arc. Arcs are the difference between two points. The equation look as follows:

$$\Delta d^{i,j}_{t_k t_0} = (1, -1) \begin{pmatrix} d^{i}_{t_k t_0} \\ d^{j}_{t_k t_0} \end{pmatrix} \tag{3.51}$$

Eq. (3.51) computes the difference between point P_i and P_j at a particular epoch. The displacements are relative w.r.t. the first epoch. Therefore, each displacement value is from epoch t_k relative to the first epoch t_0.

3.4.3. Arc Characterization

The next step is the characterization of an arc using a particular model. With the model, we can predict, to a certain degree, the displacement of the new incoming epoch. Therefore, we prefer a model describing the displacement time series to the highest degree such that the model incorporates all the expected behavior, and only the anomalous behavior will be the culprit for deviating trends.

However, for this thesis, to keep the methodology initially lean and straightforward, we decided only to use a linear model. We compute the linear model using linear least-squares (Heath, 2018). The linear least-squares only computes a displacement rate (or velocity) and a bias and neglects seasonal trends or higher polynomial trends.

With this model, we can predict whether a newly acquired observation behaves anomalously. We perform the prediction by first computing the variance of the residuals and then setting the confidence intervals. The variance computation will provide a proxy for how noisy an arc is behaving. Another aspect of this model is how we would update it with newly incoming observations. It is advantageous to incorporate newly obtained epochs since it will improve the model. However, this might also taint the model so that anomalous behavior is not detected. in the next section.

Arc Noise and Confidence Interval Determination

We describe the noise, or variance, of the displacement time series in Section. 3.2.3. Here, the same methodology is applied to describe the behavior of the arc.

The next step is to define confidence intervals. We use these confidence intervals to test if we need to flag the new incoming observation. In this study, we set the confidence interval to 95 percent. We calculated these confidence intervals by multiplying the standard deviation (obtained from the analysis of the residuals, Section. 3.2.3) by a factor $r_{\alpha/2}$. Thus, α is the significance level, and for a 95 percent confidence interval, the significance level must be 5 percent ($\alpha = 1 - 0.95$).

We assume that newly incoming measurements will fall within the defined confidence interval and thus outside the critical region (K). We can set the size of the critical region by setting the significance level. The test is two-tailed, meaning that observations above and below the mean can fall within the critical region (K). We split the critical region symmetrically into two regions with a similar size (i.e., two regions with 2.5 percent). To find this boundary, we need to compute k_α. We find k_α by multiplying the critical value ($r_{\alpha/2}$) with the standard deviation. Since the mean is non-zero, we add or subtract the multiplication from the mean (Teunissen, 2000). Using either lookup tables (Teunissen, 2000) or the SciPy library in Python (Virtanen et al., 2020), we find the critical value to be $r_{\alpha/2} = 1.96$ for a confidence interval of 95 percent (Teunissen, 2000). We then apply the obtained critical value to compute the confidence interval, and we show an example in Fig. 3.8.

Relative Days Computation

The model can not use fixed Gregorian dates. These are, however, used in the dataset to indicate the day it was measured. To compute the model, we convert the epoch dates to relative days. We set the first day of the displacement time series to zero, and every subsequent date will be in relative days since the first epoch. We automated this process with the a Python function (see Appendix A).

3.4.4. Arc Flagging Procedure

We start the arc anomaly detection procedure using one of the predefined model updating strategies. With the linear model, the arc noise is estimated (Section. 3.2.3), and the 95% confidence intervals are defined. The prediction for a new incoming observation is estimated using

$$\Delta\hat{d}^{ij}_{t_k t_0} = \hat{x}_1 t_k + \hat{x}_2. \tag{3.52}$$

Where $\hat{x}_1$ and $\hat{x}_2$ are the model estimations (velocity and bias) and t_k is the relative number of days. The confidence interval is defined by

$$|\Delta d^{ij}_{t_k t_0}| \leq \underline{\hat{y}} + r_{\alpha/2}\sigma_{\hat{e}}. \tag{3.53}$$

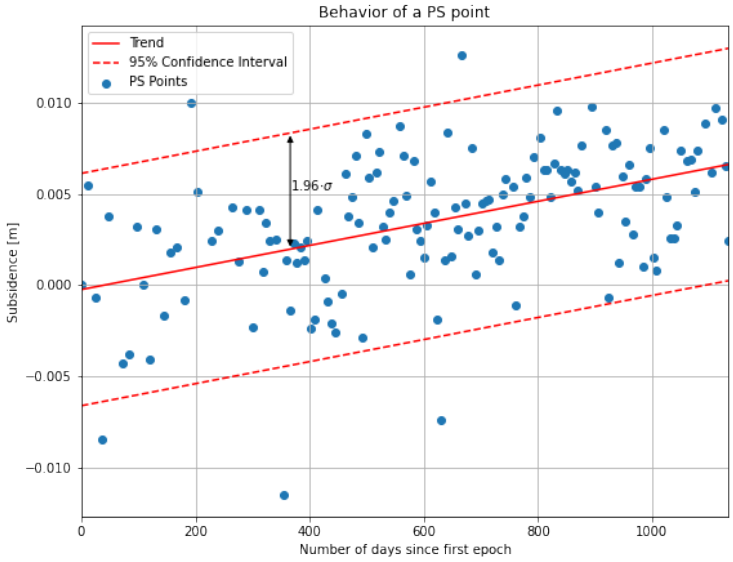


Figure 3.8: Time Series of the LOS Displacements plotted together with the trend and the subsequent 95% confidence intervals

Whereby y is the arc displacement observation, $\sigma_{\hat{e}}$ is the standard deviation of the residuals of the linear model and $r_{\alpha/2}$ the critical value and thus selecting 95% confidence interval. Eq. (3.53) tests new incoming observation.

If the new observation does not satisfy Eq. (3.53) (i.e., falls outside the boundary), we flag that particular arc. We give flagged arcs a value of 1, and non-flagged arcs will always obtain 0.

Flag count

Here, we introduce a counting method that implements a short-term memory in the flagging procedure. The flag count is the value given to a specific arc on a specific epoch. The flag count typically is zero (arc is not flagged). We increase the flag count with one when the arc is flagged. Increasing the flag count means that the flag count of an arc can obtain a very high number. We reset the flag count to zero again when the arc is not flagged. In practice, we see the flag count increase a few times; then, the method does not flag the arc, resetting the flag count to zero. We can set a maximum value for the flag count to ensure a maximum short-term memory length.

We also use the flag count to color the arcs in a plot. In this way, it is an easy method to show the arcs behaving most anomalous visually.

3.4.5. Total Flagged Value

We discussed a method that creates arcs, characterizes them, and finally flags anomalous behavior. In this section, we convert the flagged arcs to a value linked to the observation points. We named this value the *total flagged value*. The total flagged value indicates how many flagged arcs are going towards an observation point.

We compute the total flagged value by going through the tuple list defined in Section 3.4.1. Then, for each point number in the tuple, we add the flag count of a particular epoch of that tuple (i.e., arc) to the total flagged value of the selected point in the tuple. Therefore, we add the flag count of a flagged arc to both endpoints of the arc.

3.4.6. Anomalous Point Detection

The last step we need to implement is defining a threshold. Section 3.4.5 describes the total flagged value, which is an indication of how many flagged arcs and how many high flag counts are running towards a specific point. We want to flag a point as behaving anomalously. To do this, we suggest implementing a threshold value.

The threshold is a minimum value for the total flagged value above which we determine a point to behave anomalously. The threshold, however, depends on multiple factors. A few of these factors are the quality of the InSAR dataset, the local geology, the significance level, and the detectability power. Therefore, the anomalous point detection method requires tweaking and testing before implementing it as a monitoring system.

4

Results and Discussion

We present the results obtained by following the methodology explained in Chapter 3. Each of the sections will be answering one of the sub-questions. We start by showing how we modeled the impending sinkhole surface expression in Section 4.1.1. Section 4.2 will show the results how we characterized the displacement time series. Section 4.3 will show the results of the spatio-temporal strategy based on the kinematic model. We end this chapter with Section 4.4, the arc-based temporal strategy.

4.1. Sinkhole Surface Expression and Simulation

The first results will show how we simulated the surface motions in the vertical and the horizontal direction. Second, we compare the Eulerian-based method to the Lagrangian-based method. The last step is to apply our defined kinematic model to some real-world cases.

4.1.1. Surface Displacement Curves

In Section 2.2 we explained the mathematical approximation of the displacement curves. Here, we

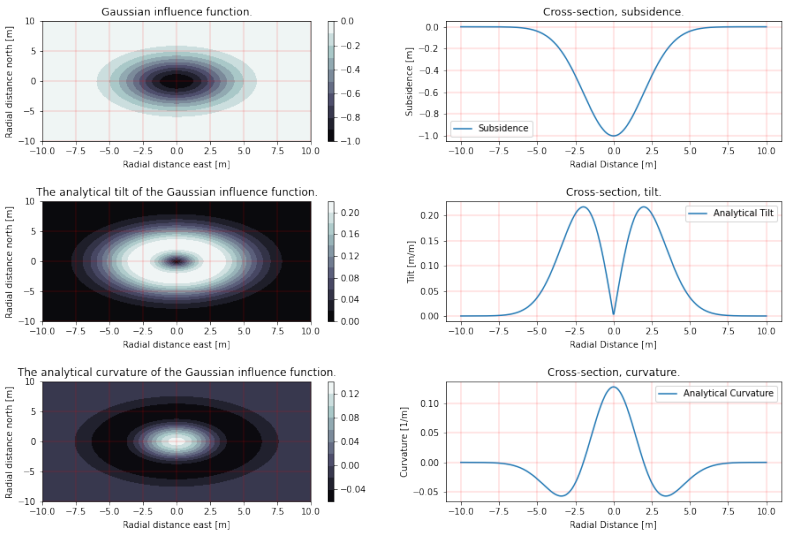


Figure 4.1: The first column shows a top view of the different curves generated by the geometric model. The second column is a cross-sectional view of the generated curves. The first row shows the subsidence curve, vertical displacement. Then, the second row shows the tilt (or slope) of the subsidence curve. The third and last row shows the curvature computed using both approaches. The radius of influence used is 7 meters.

used the geometric model described in Eq. (3.4) to compute the subsidence experience at an arbitrary epoch. We can derive other curve properties analytically, such as tilt (i.e., slope) and curvature, from the geometrical model. Fig. 4.1 shows the different curves associated with ground subsidence. We base the curves on the geometric model (Eq. (3.4)). The first column shows a top view of the type of curve in question. The second column is a cross-sectional view. The set of figures hosts the subsidence curve on the first row with the tilt on the second row and the curvature on the last row. The *analytical* tilt and curvature are the first and second derivatives of the influence function (Eq. (3.4)), respectively. We compute the tilt using Eq. (2.14) and the curvature with Eq. (2.15).

We compute the horizontal displacement using the focal point method described in Section 2.2.3. We based the focal point method upon the geometric model, and it assumes a cavity depth. The horizontal displacement is orientated radially towards the center of the sinkhole. Therefore, taking a cross-section is concentrically symmetric. We then use the horizontal displacement to compute the linear change (strain). The linear change is the change in point spacing. Depending on the horizontal displacement, the point spacing increases or decreases. We obtain the extension or compression by dividing the point spacing with a fixed unit distance (e.g., 1 meter). The extension is positive, while the compression is a negative value. These metrics can be beneficial since maximum strain indicates where earth fissures might occur and damage buildings (Kratzsch, 1983). Fig. 4.2 shows the horizon-

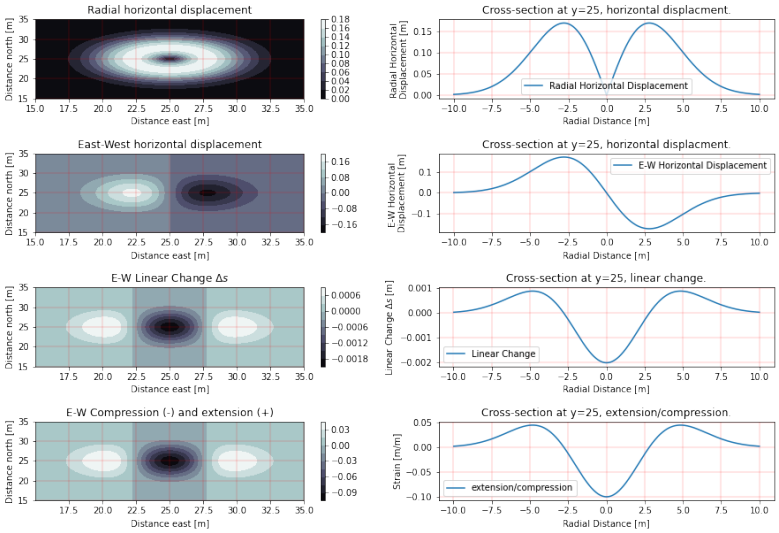


Figure 4.2: The first column is a top view of the horizontal movements. The second column is a cross-sectional view of the horizontal movements at latitude $y = 25$ meters. The first row is the radial horizontal displacement towards the center of the impending sinkhole (i.e., centripetal displacement). The second row also shows the horizontal displacement but now from East to West mimicking a descending satellite geometry. Hence, the negative horizontal displacement. The third row is the linear change (Eq. (2.21)) or change in point spacing. The last row is the extension and compression (Eq. (2.22)). The radius of influence is 7 meters.

tal displacement based on the focal point method. The first row shows the horizontal displacement orientated radially towards the center. The second row also shows the horizontal displacement but then from East to West. The third row is the linear change and the last row is the compression and extension curve.

In the next step, we begin to compare the temporal behavior of the displacement curves. We look at two types of subsidence behavior, linear and accelerating subsidence (Fig. 4.3). Here, we simulated displacements, and thus, the displacements have no real-world meaning. Also, comparing displacement values between the two figures does not bear any significance. The important message of Fig. 4.3 is the temporal ground movement. Fig 4.3 shows the linear subsidence and horizontal displacement with a linear subsidence velocity. Each extra time step produces an equal amount of vertical and horizontal displacement.

Comparing this behavior to the accelerating subsidence, we see that the accelerating subsidence

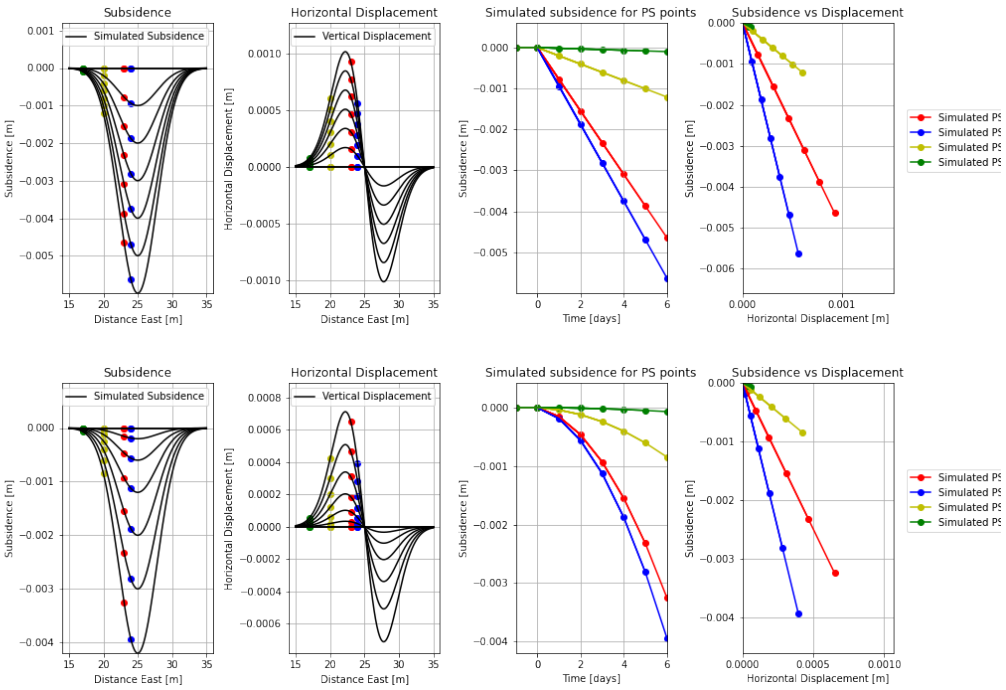


Figure 4.3: The top figure shows the linear subsidence behavior, and the bottom figure shows an accelerating subsidence behavior. Each column contains the same type of displacement. The first column shows the vertical displacement. The increasing subsidence shows how an upward moving cavity would interact with the surface. The second column is the vertical displacement. In both the first and second columns, we mark four points on the curve. The third column shows these marked points and how they temporally move vertically. The last column shows the vertical displacement (y-axis) and horizontal displacement (x-axis).

starts with minimal subsidence but quickly increases. We observe the same behavior in the horizontal displacement figure. In the first two columns, we selected four points which we highlight in the next two columns. In the third column, we see the temporal vertical movement of a single point. Depending on the location in the subsidence trough, the point moves downward. The last column shows the vertical versus the horizontal displacement. The vertical displacement is dominant, but the horizontal displacement can still get to 50% of the vertical displacement.

4.1.2. Eulerian versus Lagrangian Displacement Comparison

In this section, we compare the two viewpoints. We base the Eulerian viewpoint on the Gaussian kinematic model and base the Lagrangian viewpoint on the Mogi source model (physical model). Thus, we compare the subsidence curves obtained from the kinematic and the physical model. We adjust the geometric model to fit the physical model. We can manipulate the geometric model by modifying R or $S_{\max}$. We set R equal to 7 meters and $S_{\max}$ to the maximum subsidence obtained from the physical model. The physical model has more parameters to tune. First, we compare the vertical displacement

Table 4.1: Values for the parameters used for the physical model.

Parameter	Value	Source
ΔV	0.5 m^3	Eq. (3.10)
ν	0.25	Bekendam & Pottgens (1995)
x_0	0	-
y_0	0	-
z_0	-11.5 m	$d = H = \frac{R}{\tan(\theta)}$
θ	35^o	Ren & Li (2008)

curves of the two models. We used the maximum subsidence to reference both curves. We assume that the volume decreases, resembling a sphere with a one-meter diameter (creating a volume loss of

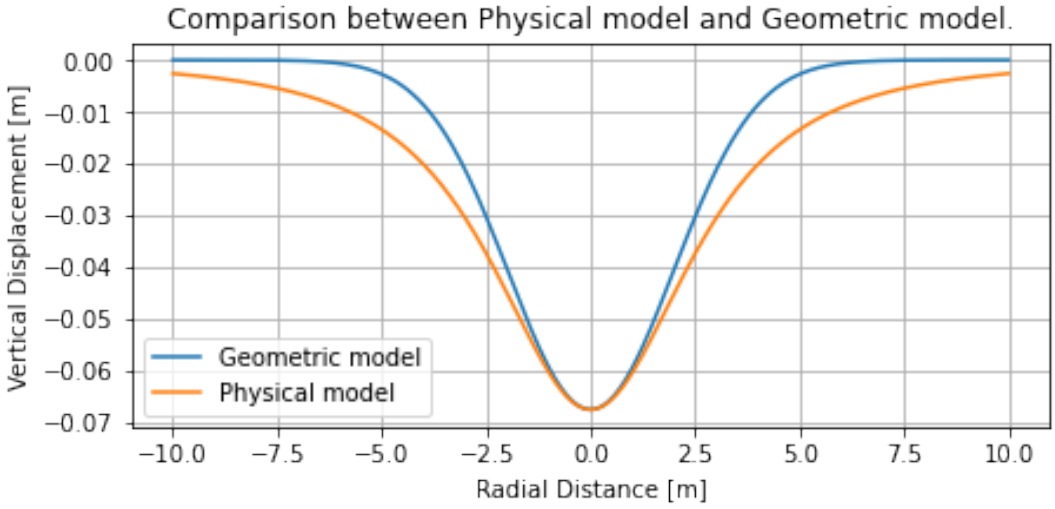


Figure 4.4: The two curves represent the two models used in this section. The figure shows the vertical displacement. The blue curve is the geometric model and the orange curve is the physical model. The radius of influence for both models is set to 7 meters.

around 0.5m^3). The Mogi model will generate maximum subsidence straight above the cavity location. We then use the maximum subsidence to adjust the geometric model. Both models now have the maximum subsidence in common, and we can now look at how the models express an impending sinkhole.

Fig. 4.4 shows the curve for the geometric and the physical model. We see that the maximum subsidence coincides with both curves. The geometric model has set the radius of influence to seven meters. The equation of the geometric model explicitly uses the radius of influence, while the Mogi model implicitly uses the radius of influence. We compute the depth of the cavity (d) using the radius of influence. The geometric model matches the radius of influence satisfyingly, while the Mogi model does not. The Mogi model still shows roughly a tenth of the total subsidence at the edge of the radius of influence. Therefore, concluding that the sinkhole based on the Mogi model would have a much larger surface expression.

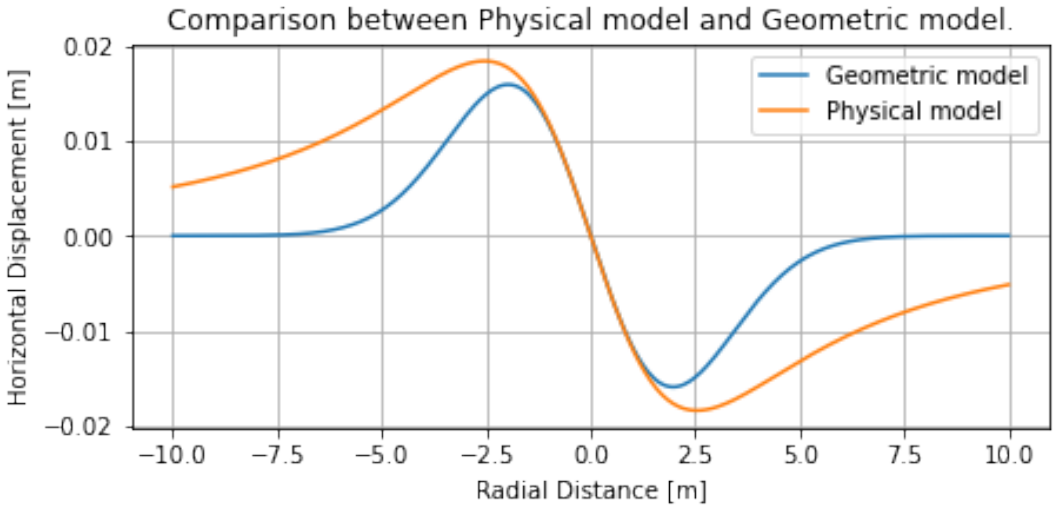


Figure 4.5: The two curves represent the two models used in this section. The figure shows the horizontal displacement. The blue curve is the geometric model. The horizontal component for the geometric model is obtained with the focal point method (see Section 2.2.3). The orange curve is the physical model. The radius of influence for both models is set to 7 meters.

Fig. 4.5 shows the horizontal displacement obtained using the geometric model and the physical

model. We compute the horizontal displacement for the geometric model using equations described in Section 2.2.3 (focal point method).

We notice that both curves behave similarly near the center of the sinkhole. The curves start to deviate nearing the peak displacement. The peak of the geometric model is near 2 meters from the center, while the maximum displacement for the physical model is closer to 2.5 meters from the center. After the maximum displacement, the geometric model converges to zero and is zero at the 7-meter border. Yet, the physical model still exhibits more than half of its maximum displacement. The large magnitude of the horizontal displacement suggests that there is more horizontal displacement present than expected initially.

Next, we take a closer look at how the models behave over time. We implement this by using the physical model. The physical model reacts to volume change and repositions the affected points. Therefore, we take an arbitrary amount of epochs (in our case, 10) with an arbitrary unit (months, days, hours, minutes). At each epoch, the same amount of volume decreases, affecting the tracked points. At each epoch, the physical model computes a new curve. We compute the new curve by taking the z coordinates of the points and subtracting the computed vertical displacement. We couple the geometric model at each epoch to the physical model using the maximum subsidence. The physical

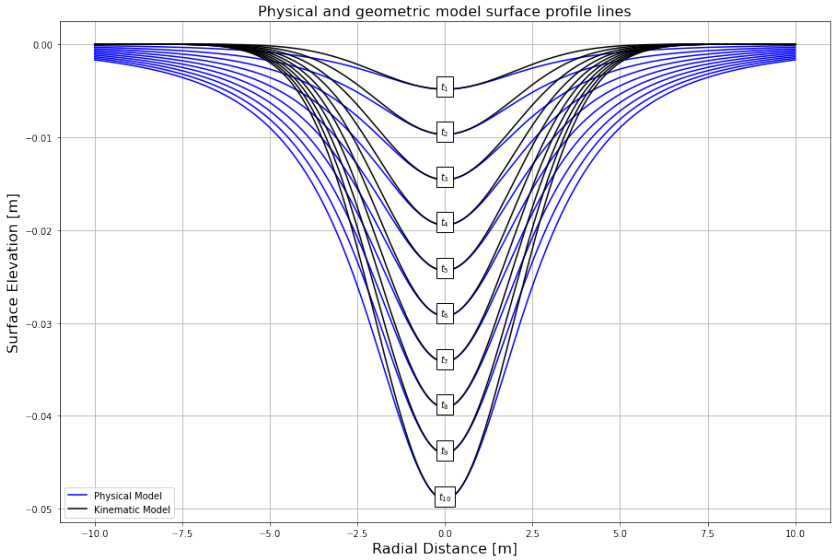


Figure 4.6: The subsidence patterns computed by the geometric and the physical model. At each new time step, both models compute the new shape of the subsidence curve.

model computes the new position of each point in 3D space (x, y, z). Therefore, the further in time we are, the more the points have moved relative to their starting position. Fig. 4.6 shows the subsidence curves at each epoch t_i. At t_0, both models have a subsidence curve of only zero (physically meaning no subsidence). Through time, the subsidence increases, and the subsidence trough deepens. We can couple the deepening trough to the decreasing ground elevation indicative of an impending sinkhole.

We have linked the maximum subsidence to the volume decrease. At each new epoch, the same amount of volume shrinkage occurs. Thus, when we plot the maximum subsidences, we expect a linearly decreasing line. Fig. 4.7 shows the decreasing trend of the maximum subsidence. However, when we increase the volume change per epoch, the trend starts to exhibit accelerating behavior.

The right figure in Fig. 4.7 shows the accelerating trend. We increased the volume decrease per epoch to unrealistic high values, but we implemented this absurd decrease to observe the accelerating trend better. Thus, when the surface is getting close to the cavity, the subsidence process will accelerate. We made this conclusion based on the physical model.

We do the same for the horizontal displacement. Only here we show the displacement relative to t_0, not the actual displacement values. The reason is to be able to compare both models. In Fig. 4.8, we see that close to the center, both models behave in the same way. However, the differences are starting to appear going towards the peaks of the curves. In this case, the peak of the physical model is later (further from the center) and higher than the geometric model.

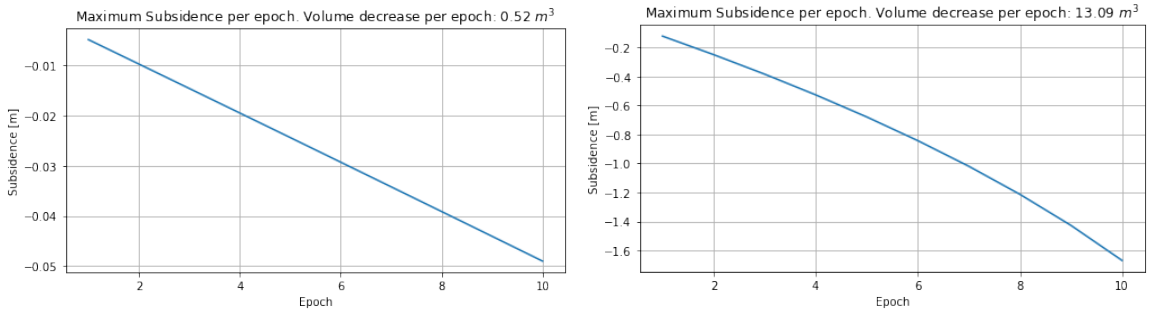


Figure 4.7: Maximum Subsidence of the Mogi model with to different volume changes. The left figure shows the temporal change of the maximum subsidence. The right figure also shows the temporal change of the maximum subsidence, but here the volume change is 25 times higher.

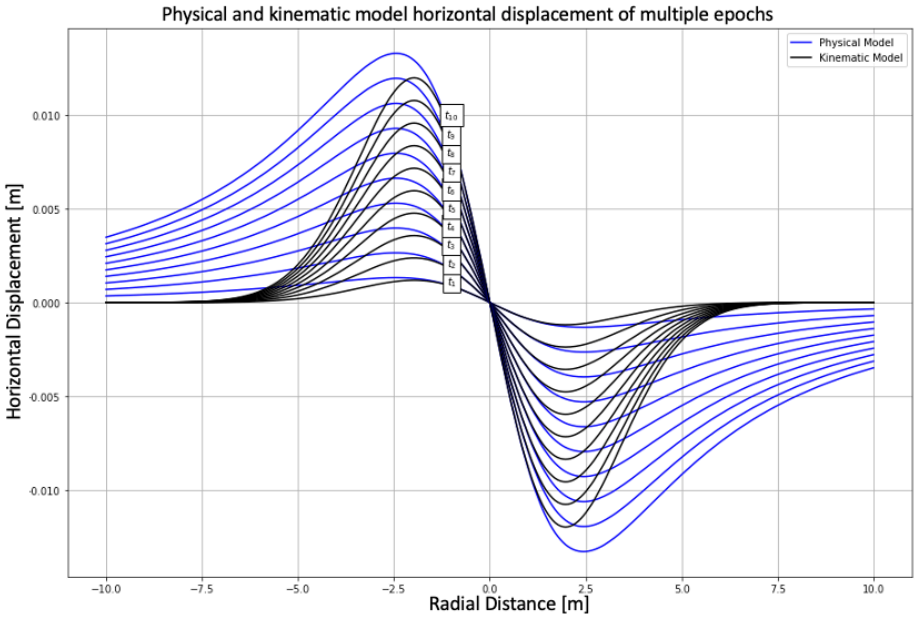


Figure 4.8: The horizontal displacement of the geometric and the physical model.

When we increase the volume change per epoch, we notice an interesting effect in the curves produced by the physical model. Fig. 4.9 shows the horizontal displacement with the maximum dis-

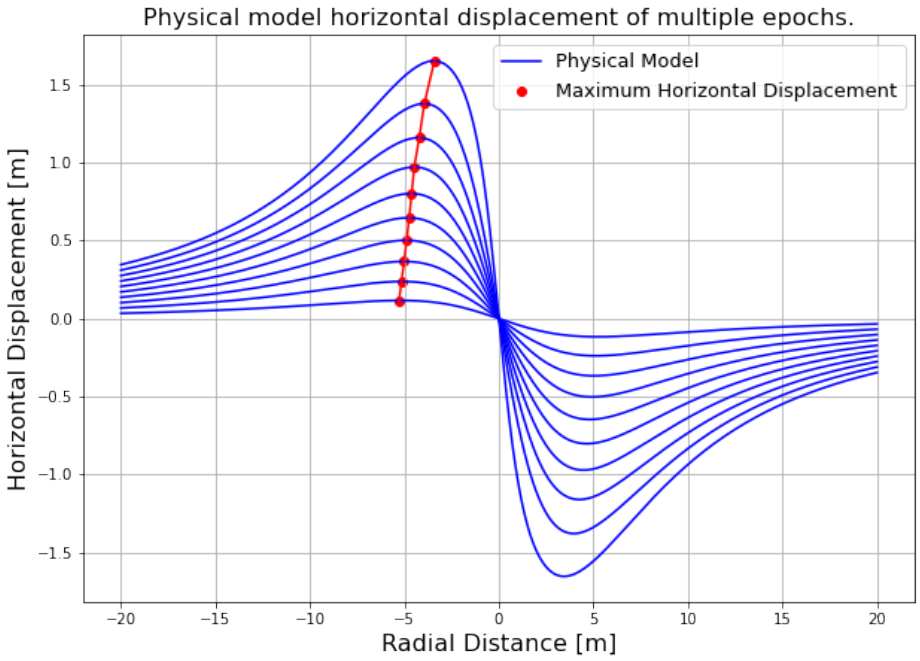


Figure 4.9: The horizontal displacement computed using the Mogi model. The volume change is increased to extremes exacerbating the shift in postion of the peaks.

placement marked in red. The point of maximum displacement shifts slowly towards the center. We increased the volume change just as in Fig. 4.7 to observe the behavior better. This moving trend is also present in Fig. 4.8, but not significant enough. Hence, we can conclude that this effect is present but not significant in the simulation of a sinkhole.

The vertical displacement is maximum in the center and slowly levels off radially outward. The horizontal displacement is zero in the center, increases to a peak, and then levels off towards zero. However, the vertical and horizontal displacement levels off at different rates. The difference is significant enough that the vertically dominant displacement becomes smaller than the horizontal displacement at some point. Fig. 4.10 shows when the displacements in the models are vertically dominant and

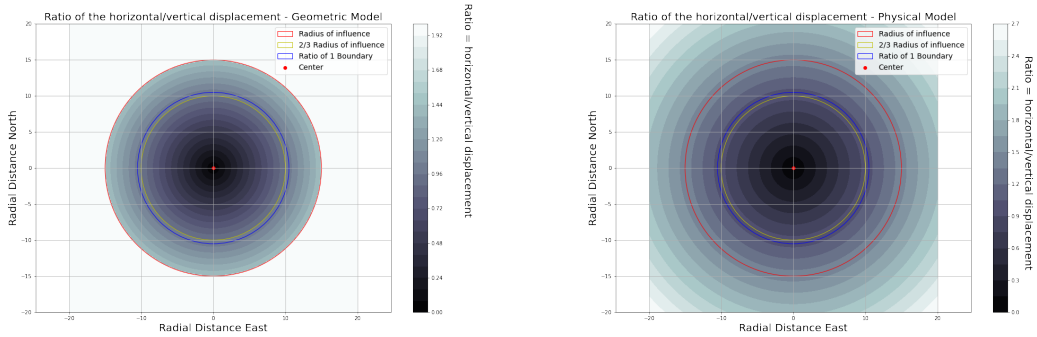


Figure 4.10: Here, we show the ratio of the horizontal versus vertical displacement direction. Close to the impending sinkhole center, we observe a dominant vertical displacement, while at the edge, the horizontal displacement is dominant. We base the left figure on the geometric model and the right figure on the physical model. The theory for the geometric model states that there is no significant displacement outside the radius of influence. Hence, no displacement.

horizontally dominant. Both models display a transition at around the same point within the predefined radius of influence. The shift is roughly around $\frac{2}{3}$ of the radius of influence. Thus, the models suggest that the radius of influence is larger than the possible observed depression and that the horizontal displacement becomes more relevant the further away from the center.

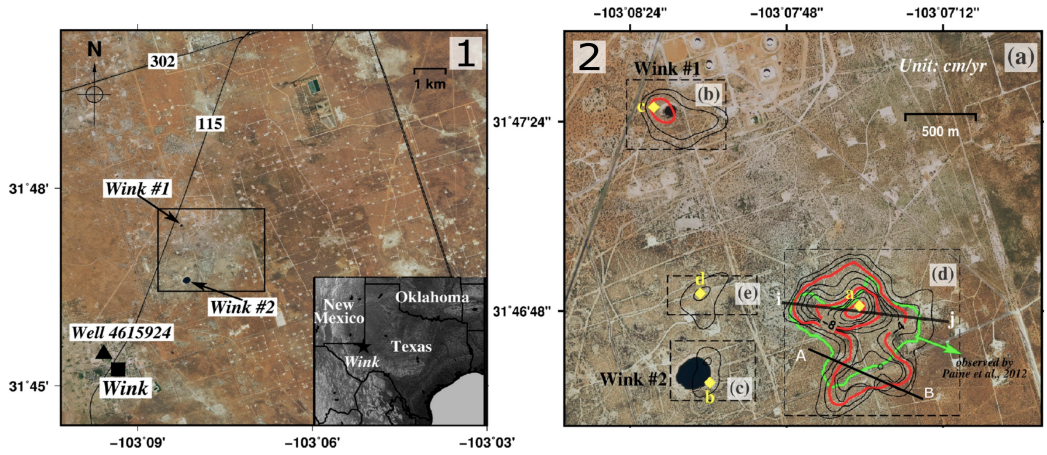


Figure 4.11: Location of the Wink sinkholes and the newly forming unstable locations. Subfigure (1) shows the location where Wink is located in Texas. Subfigure (2) shows Wink Sink #1, #2, and the unstable area. Line i-j is the cross-sectional location found in Kim et al. (2016) and line A-B is the cross-sectional location found in Kim et al. (2019). Both images are from the paper Kim et al. (2016). Both subfigures are adjusted with a number for better reference. In subfigure (2), line A-B has been added.

4.1.3. Comparing the kinematic model to real-world cases

This section applies different models (physical and kinematic) and tests them to real-world cases found in three separate published papers. The selected cases are from two separate areas. The first area is Wink, located in Texas, USA, and the second area is the west coast of the Dead Sea in Israel. This comparison aims to test whether the defined models are flexible and adequately describing the observed displacements related to sinkhole subsidence (describe at least 80% of the subsidence, see Section 3.1.6).

We explained the error metric in Section 3.1.6 on which we test, characterize, and evaluate the correctness and the behavior of the used models. We provide some (geological) background information together with the initial displacement time series for each location. In the last step, we compare the selected models to the digitized real-world displacement time series.

We fit the model to the data using the linear least-squares method and compute two parameters, v, and R. The model is linearized to compute the two parameters. We describe the implementation of the linear least-squares in Section 3.1.4.

Wink Texas, USA

Sinkholes in West Texas have developed from the dissolution of subsurface evaporite deposits that came in contact with fresh groundwater. The first sinkhole was formed on 3 June 1980 (Wink Sink #1). The second sinkhole (Wink Sink #2) was formed on 21 May 2002 and is located about 1500 meters south of Wink Sink #1. The petroleum activity from 1926 to 1964 around the Wink sinkholes has been suspected to be a trigger that accelerated the dissolution of the underlying salt bed (Johnson, 1989). The cavity then grew and migrated upward and resulted in the formation of the Wink sinkholes after successive roof failures. The region around the collapsed sinkholes in Wink is still subsiding, and the neighboring areas of Wink Sink #2 are significantly unstable (Kim et al., 2019; Paine et al., 2012). Fig. 4.11 shows the locations of the collapsed sinkholes and a newly developing unstable area. The papers Kim et al. (2016) and Kim et al. (2019) each took a closer look at a part of the unstable developing area. Fig. 4.11-2 shows two lines, i-j and A-B. These are both cross-sectional lines. The reason for the odd numbering of the cross-sections is that we preserved the papers' notation. The displacements measured at both cross-sections in Fig. 4.11-2 are shown in Fig. 4.12. The vertical deformation is the most dominant component of the ground deformation around Wink sinkholes and

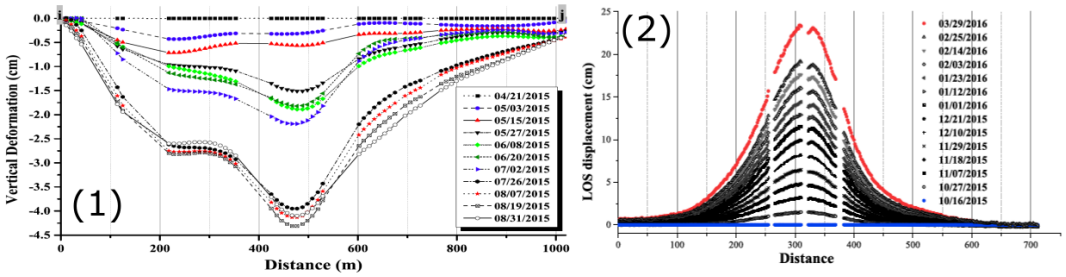


Figure 4.12: The InSAR measured displacements at Wink. Subfigure (1) is the vertical subsidence of line i-j in Fig. 4.11-2 obtained from Kim et al. (2016) and digitized. Subfigure (2) is the line of sight displacement of line A-B in Fig. 4.11-2 obtained from Kim et al. (2019) and digitized.

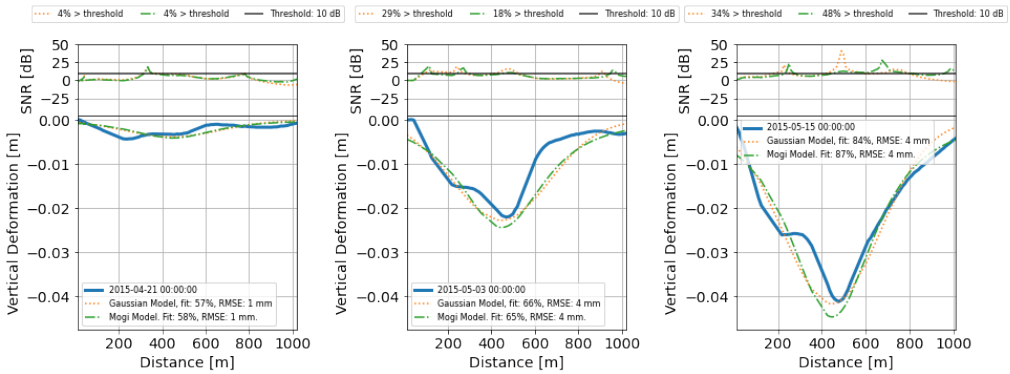


Figure 4.13: Fitted kinematic and physical models to digitized sinkhole data obtained from Kim et al. (2016).

can be accounted for 80% of the LOS InSAR measurements (Kim et al., 2016). Therefore, in the analysis, the LOS displacements in Fig. 4.12-2 are assumed to be very similar to the vertical subsidence. The two subfigures in Fig. 4.12 are different cross-sections of the same unstable area published in two different papers, and the obtained data comes from two different satellites. Therefore the cross-section i-j in Fig. 4.11-2 is referred to as *Wink 2016* and the cross-section denoted as A-B is referred to as *Wink 2019*.

We have introduced the various models. Here, we finally implemented the models to start the analysis. The first figure is from *Wink 2016*. Fig. 4.13 shows the cross-section of *Wink 2016* with two fitted curves in the bottom part and the signal-to-noise ratio in the upper part together with the threshold. On top of the plots are the percentage values for the part of the curve above the threshold. Table 4.2 shows the average results for the RMSE and the fit for the selected models. The second

Table 4.2: Table for the averaged estimated parameters of the final results for the Wink 2016 sinkhole.

Model name	Average RMSE (mm)	Average Fit (%)
Geometric Model	2.7	74
Physical Model	2.5	77

cross-section *Wink 2019* has a very clear subsidence pattern and Fig. 4.12-2 shows the location. When we look at Fig. 4.14, we see an excellent estimate of the models. The geometric model looks like an almost perfect fit to the subsidence pattern measured. The linear subsidence rate of the models is an excellent fit to both the sinkhole cross-section of *Wink 2016* and *Wink 2019*. The SNR plots show a good correlation between maximum subsidence and the center of the sinkhole models, meaning that the most hazardous location is well estimated and can be observed very early in the sinkhole formation process. Looking at table 4.2 and 4.3 we see that all the models have a rough fit of 80% and thereby indicating that the models are well suited.

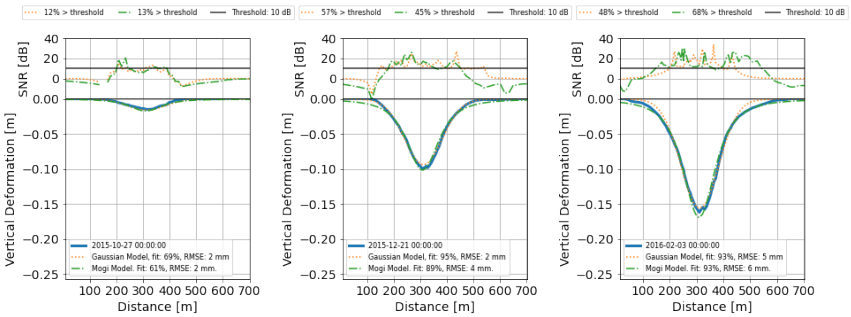


Figure 4.14: Fitted kinematic and physical models to digitized sinkhole data obtained from Kim et al. (2019).

Table 4.3: Table for the averaged estimated parameters of the final results for the Wink 2019 sinkhole.

Model name	Average RMSE (mm)	Average Fit (%)
Geometric Model	3.9	90
Physical Model	4.5	87

Dead Sea, Israel

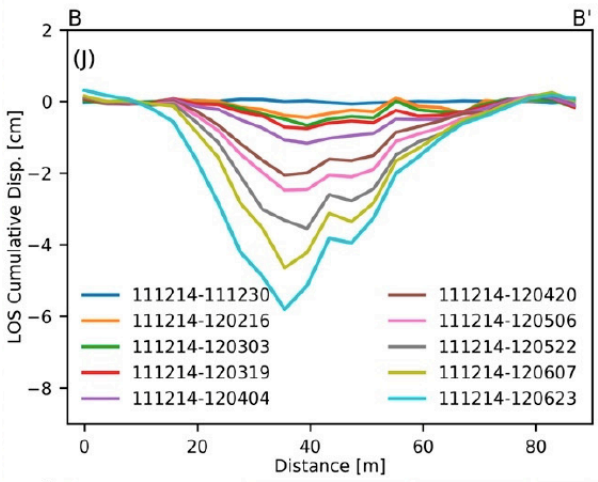


Figure 4.15: Line of sight cumulative displacement time series from the published paper Nof et al. (2019).

More than 6000 sinkholes have been mapped in the region of the Dead Sea (Frumkin & Raz, 2001). The hazard of sinkholes affects the region's local agriculture, industry, tourism, infrastructure, and daily life (Nof et al., 2019). The Dead Sea has a 5-20 meter thick salt (halite) layer buried at depths between 5 m to 65 m along the shores. The water level has been declining at a rate of more than 1m/yr. The result is the exposure of the salt layer to unsaturated water, and the formation of cavities eventually leads to collapses in the shape of sinkholes (Nof et al., 2019). Much research has already been performed on the sinkhole phenomenon at the Dead Sea. Also, the use of InSAR has already been researched for some time. In the paper Nof et al. (2019) the Hever site has been a case study where a sinkhole collapse has been investigated using InSAR. The displacement time series obtained from the ex post facto analysis was used to apply our models. Fig. 4.15 shows the los cumulative displacements measured at the Hever site. The time windows starts on 14 December 2011 and closes on 23 June 2012. The time window is roughly half a year. The displacement is in the line of sight and shows a maximum displacement of nearly 6 centimeters. Here, we also assumed that displacements in the line of sight are sufficient to test our subsidence models. Fig. 4.16 shows the two models with their best estimate. The estimate seems to overperform in the first two figures and underperforms in the last figure. The bad performance in Fig. 4.16 is also visible in Table 4.4. The average RMSE is high, but the average fit is close to zero. The linear subsidence approximation of this sinkhole is thus not adequate for all the tested models. We then try to use an accelerating approximation for our models.

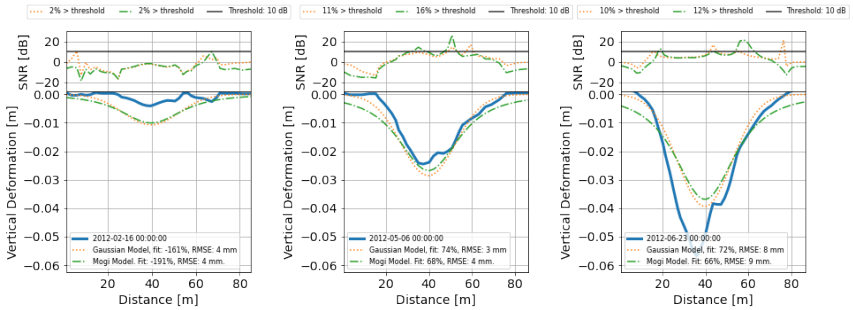


Figure 4.16: Fitted linear kinematic and physical models to digitized sinkhole data obtained from Nof et al. (2019)

Table 4.4: Table for the estimated parameters for the linear subsidence approximation shown in Fig. 4.16. The Dead Sea sinkhole is found in the paper Nof et al. (2019).

Model name	Average RMSE (mm)	Average Fit (%)
Gaussian	4.3	3.5
Mogi	4.8	−13

We describe the accelerating kinematic model in Eq. (3.8). Eq. (3.8) improves the performance of the

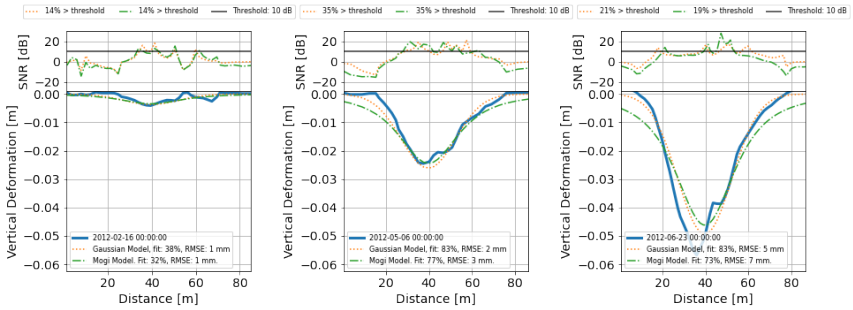


Figure 4.17: Fitted accelerating kinematic and physical models to digitized sinkhole data obtained from Nof et al. (2019).

kinematic model and we show the results in Fig. 4.17. We show the averaged results in Table 4.5. The average RMSE has improved satisfactorily, but the average fit still underperforms. The last few epochs (in Fig. 4.17) show a terrible estimate, and an explanation can be that the sinkhole experienced partial collapse at this stage.

4.1.4. Discussion

We discern from Fig. 4.1 that the obtained tilt and curvature from the computed vertical displacement differ in magnitude for the numeral and the mathematical approach. The shape shows a likeness since the peaks and troughs of both methods share the same radial position.

Fig. 4.2 shows the horizontal displacement. The horizontal displacement assumes a cavity depth to compute the displacements. The cavity depth is related to the angle of draw θ and the radius of influence (R), Eq. (2.4). R governs the surface expression, and θ is characteristic for the local geology. Therefore, θ can be roughly indicative of the local geology (Ren & Li, 2008). Both parameters might be hard to determine in practice since the surface expression might be faint or the local geology is challenging. The depth estimation of the cavity might not be reliable.

The horizontal movement is radially towards the center of the sinkhole. We see that the maximum horizontal displacement is roughly at $\frac{1}{3}$ of the R from the center. The linear change shows how each point on the surface gets dislocated w.r.t. its neighbor. The linear change is also active beyond R. When we divide the linear change by a fixed unit distance (e.g., 1 meter), we get the extension and compression. We see that around roughly $\frac{2}{3}$ of R we have the maximum extension. The position of maximum extension is the location where the first signs of cracks will appear (Kratzsch, 1983). At

Table 4.5: Table for the averaged estimated parameters for the accelerating subsidence approximation. The final results are from the Dead Sea sinkhole found in Nof et al. (2019).

Model name	Average RMSE (mm)	Average Fit (%)
Gaussian	2.1	64
Mogi	2.9	57

around $\frac{1}{3}$ of R, we see that the extension flips to compression (positive to negative). At the center of the sinkhole, the compression is about twice as big as the extension at its peak.

Fig. 4.3 shows the temporal behavior of the vertical and horizontal displacements. The centers subside at a linear and accelerating rate according to our implementation. We notice that the peaks stay at the same position. The vertical and horizontal displacements increase as expected, and we observe the linear or accelerating trend in the third column. The last column shows the vertical versus the horizontal displacement.

We notice from the last column that closer to the middle, the vertical displacement is more dominant, while closer to the edges, the horizontal displacement becomes more dominant. It follows roughly the following rule of thumb; the horizontal displacement is $\sim 10\%$ of the vertical displacement in the center. Closer to the edges, the ratio approaches $\sim 50\%$.

In Section 4.1.2 we started by explaining the Eulerian-based and the Lagrangian-based viewpoint. In this study, we mainly apply the Eulerian-based geometric and kinematic models. However, the Lagrangian-based model is closer to reality since it tracks individual points. The reason why we prefer using the Eulerian-based model is that it is easier to implement. Therefore, in this section, we looked at the similarities and differences between the two viewpoints. For the Eulerian viewpoint, we used a geometric model based on the Gaussian influence function. For the Lagrangian viewpoint, we used the Mogi source model.

We first compared the subsidence curves of both models. We coupled the maximum subsidence of both models to investigate the shape better. We saw that the shape agrees reasonably at the center but deviates further away from the center. The geometric model is confined within the predefined radius of influence, while the subsidence simulated by the physical model extends very far out. The implications are that the surface expression for the physical model is much larger.

Second, the comparison is between the horizontal displacements generated by both models. In Fig. 4.5 we also see that the curves of both models are very similar to each other and deviate moving outward. We also see that the physical model again doesn't confine itself to the radius of influence.

In the next step, we look at how the shape of the subsidence bowl evolves. Each epoch, we again scale the geometric model to the physical model. The same volume shrinkage occurs in each epoch, and therefore, the points change slightly for the physical model. It is not the case for the geometric model. We observe the same behavior as in the single epoch for the vertical and horizontal displacement at each epoch. Interestingly, when we increase the decreasing volume (to extreme levels), we notice accelerating maximum subsidence and an inward moving maximum for the horizontal displacement. These trends are only visible using the physical model and are due to the changing point coordinates. Less intense volume change will provide a visible linear subsidence rate and an imperceptibly slow movement of the maximum horizontal displacement. We can conclude from this finding that low volume change and small displacements (centimeters) do not significantly impact the point position to induce a visible accelerating behavior. Hence, simulating sinkhole subsidence can be approximated using linear trends.

What is an interesting observation is that the horizontal displacement is significant. The horizontal displacement gets relatively larger than the vertical displacement the further away from the center. At around $\frac{2}{3}$ of the radius of influence, the horizontal displacement becomes larger than the vertical displacement. Interestingly, this holds for both the geometric and the physical models. Therefore, it seems that both models are adequate to simulate sinkhole subsidence.

In Section 4.1.3 we see mixed results showing that we can simulate the surface expression of an impending sinkhole using one of these models. It is not perfect, and sometimes the estimate can be dubious, but it is a decent approximation in general. Relating the results to the Pareto principle (mentioned in Section 3.1.6), we can conclude that the results are very relevant to the amount of subsidence and the local geology. For example, Wink showed a very good fit using a standard linear

subsidence approximation, while the sinkhole at the Dead Sea needed an accelerating subsidence approximation. Unfortunately, the accelerating subsidence approximation still didn't provide the fit according to the Pareto principle. Also, we suspect that in the last figure in Fig. 4.17 partial collapse has occurred which has a negative impact on the model fit.

The SNR shows that the signal is indeed the strongest at the center of the of the sinkhole. However, obtaining a signal above the 10 dB is not very common. The SNR also shows that the model fit works better with higher displacement values.

4.2. Spatial and Temporal Time Series Characteristics

In this section the spatial and temporal characteristics of the data are visualized using two developed methods. The methods are the Minimum Detectable Size (MDS) and the results are shown in Section 4.2.2. The MDS immediately shows, on a map, how the distribution of point scatterers lead to the detectability of an impending sinkhole. Then, the next method will provide a probability that a measured displacement can be attributed to an impending sinkhole. This method will be discussed in Section 4.2.3.

4.2.1. Subset selection and artificial impending sinkhole Implementation

The methodology for artificially superimposing an impending sinkhole signal into a real dataset was discussed in Section 3.2.1. In the introduction, we explained that we are limiting ourselves to one

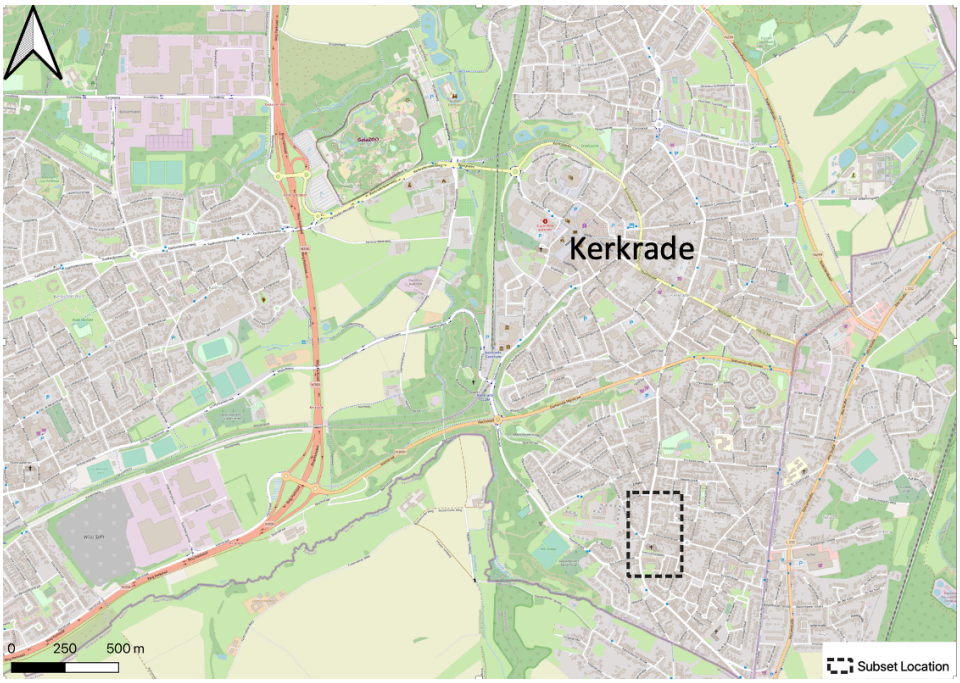


Figure 4.18: The location of the subset used for this thesis. The subset is situated in the city of Kerkrade close to the German border.

sensor and one track. From that track, we created a subset to test our methodology. We've chosen the Sentinel-1 sensor because it is free of charge. Choosing track 88 was an arbitrary decision since track 37 and track 139 would also be possible. Here, we implemented the sinkhole around the location of the Franciscanerstraat in Kerkrade.

Fig. 4.18 shows the location of the subset in Kerkrade. In this thesis, we mainly use the Rijksdriehoek coordinate reference system (EPSG:28992). It is a local Dutch reference system and has the benefit of displaying the coordinates and distances in meters. We denote the x-coordinates as *rdx* and the y-coordinates as *rdy* when we are working with Rijksdriehoek coordinate reference system. Fig. 4.19 shows the situation of the subset of the observations, or Persistent Scatterer (PS) points. The subset is originally obtained from the Sentinel-1 satellite orbiting in the descending track 88. The displacement data is originally still in the line of sight and not in the vertical direction. However, to test

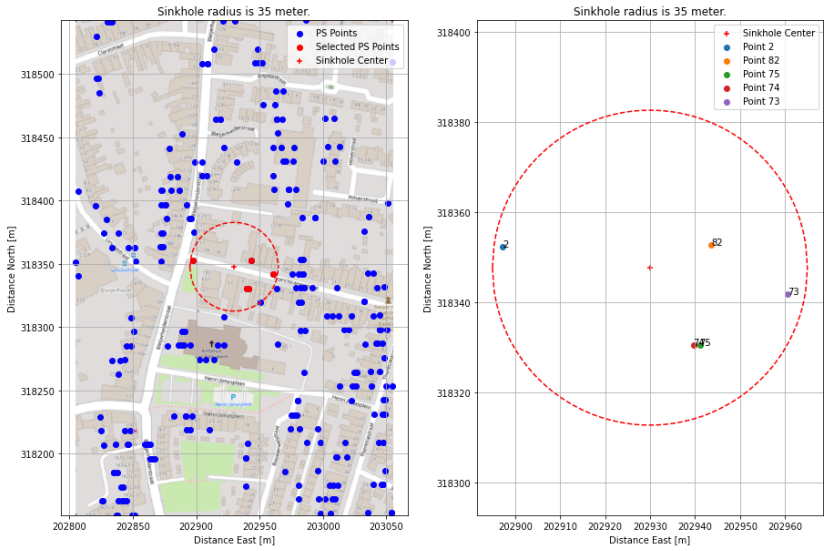


Figure 4.19: Dataset with the impending sinkhole and the affected observations. The left figure shows an overview of all the observations located within the subset. In the center the impending sinkhole center and radius of influence is shown in red. The red dots are the data points located within the radius of influence. The right figure shows the impending sinkhole center and radius of influence together with the observations located inside the radius.

the methodology, the point geometry and natural displacement variability was more important than actual vertical displacement data. The following strategical research assumes that no significant small scale, short time frame deformations are happening in the subset and that the superimposed sinkhole is the only deformation event.

The left figure shows the whole subset. In the center of the subset, the superimposed sinkhole is shown. The sinkhole center (+) and radius of influence (−−) is shown in red. The radius of influence is 35 meters and thus the diameter becomes 70 meters. This is considerable for a sinkhole, but for research purposes this sinkhole is enlarged to better investigated the signal and behavior it is generating. The red dots in the left figure show the affected observation and are also shown in the right figure. The right figure is zoomed in on the implemented sinkhole phenomenon. The observations are numbers such that the time series in later figures are more easy to compare. Also, the distance of each observation towards the center (location of maximum subsidence) is clearly visible. For example, observation 82 is the closest observation to the center and thus we can expect to observe the strongest subsidence signal in its time series.

Linear Subsidence Velocity

The location for the superimposed sinkhole is selected and now the sinkhole signal is added to the displacement time series. In this section the sinkhole with a linear subsidence velocity is added to the displacement time series. Fig. 4.20 shows two figures where the sinkhole signal is added to the displacement time series. The left figure is a top view of the observations. Each observation is color with respect to its deformation rate, or displacement velocity. The displacement velocity is computed by fitting a linear model through the displacement time series of each observation. So, in this way the general movement of an observation can be characterized. The general movement is shown as millimeter per year. The implemented sinkhole is also shown on the map. From the left figure an increasingly downward trend can be seen for the points closer towards the center of the sinkhole. This, however, doesn't necessarily mean something is actually happening at this locations (if the locations of the sinkhole was unknown).

The right side figure shows the displacement time series of all the observations within the radius of influence. Each line corresponds with a particular observation and can be compared with the observations in the right figure of Fig. 4.19. Be aware, the colors in Fig. 4.19 and Fig. 4.20 do not agree. After the vertical dotted black line, we added the impending sinkhole signal to the dataset.

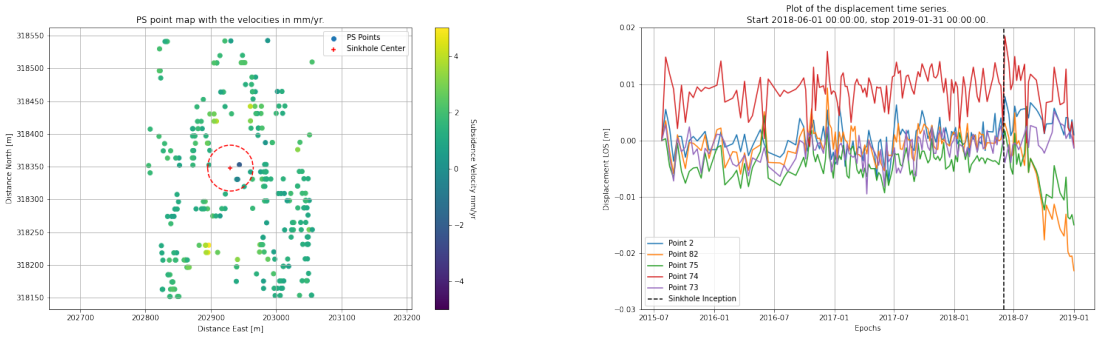


Figure 4.20: The left figure shows the observation colored by their displacement velocity. The displacement movement is displayed in mm/yr. The location of the sinkhole is also visible in the left figure. The right figure shows the displacement time series of all the observation within the sinkhole radius of influence. The vertical dotted black line shows the moment when the sinkhole signal is added to the time series.

Accelerating Subsidence Velocity

An impending sinkhole doesn't always subside linearly. Sometimes it exhibits an accelerating behavior. In this section an accelerating trend is added to the dataset. Fig. 4.21 shows in the left figure the

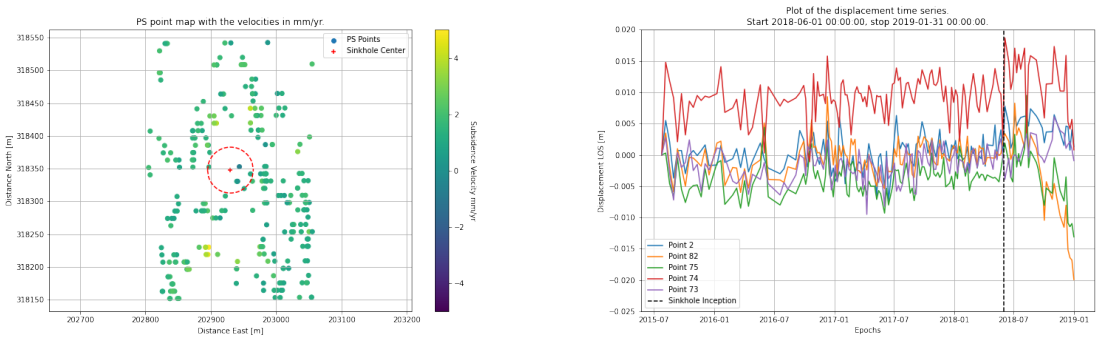


Figure 4.21: The left figure shows the observation colored by their displacement velocity. The displacement movement is displayed in mm/yr. The location of the sinkhole is also visible in the left figure. The right figure shows the displacement time series of all the observation within the sinkhole radius of influence. The vertical dotted black line shows the moment when the sinkhole signal is added to the time series.

observation with the subsidence velocity. The right figure shows the displacement time series with the accelerating sinkhole signal added to the end of the series. The moment when we added the impending sinkhole signal to the dataset is highlighted with the vertical dotted black line and continues to the end of the dataset.

4.2.2. Minimum Detectable Size

Before we start implementing one of our strategies on a particular area of interest, we prefer to have some a priori knowledge. We can obtain this knowledge in the shape of maps and show the dataset's characteristics in the area of interest. The maps are the minimum detectable sinkhole size and the minimum detectable deformation given a detectability power and significance level. The latter map can also be inverted, showing the detectability power given a minimum detectable deformation and significance level. These maps will provide us with a priori knowledge and help us manage expectations.

First, we show the map for the minimum detectable sinkhole size of our area of interest. Fig. 4.22 shows the subset of observations of the Sentinel-1 sensor in its ascending track number 88. The map clearly shows that closer to clusters of observations, the minimum size gets smaller. The smallest sinkhole tested in this dataset is a sinkhole with a radius of influence of 1 meter and thus has a diameter of 2 meters. We used a posting of 1 meter, meaning that every 1 meter we evaluated the algorithm. The edges of the map may show some misleading results because the dataset is a subset.

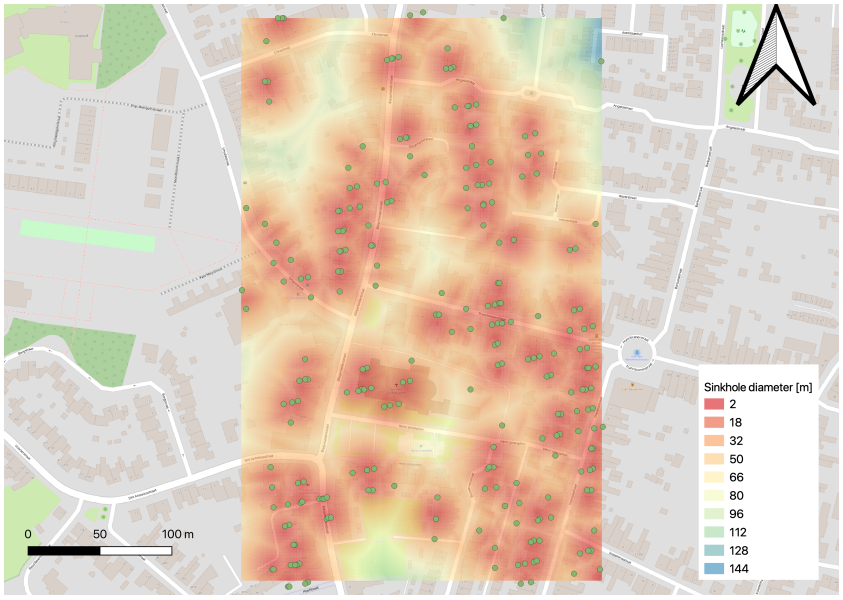


Figure 4.22: Minimal detectable size of a potential impending sinkhole over the area in Limburg around the Franciscanerstraat in Kerkrade. The observations are from the Sentinel-1 sensor and the ascending track 88.

The observations near the edges are filtered out since we are using a subset. The results of the map are interesting. It indicates what size is possible to observe and where we can observe it. We are more interested in sinkholes in urbanized areas than in nature since it poses an anthropological hazard. Fig. 4.23 shows a histogram of the observed sizes. Interestingly, minimum detectable sizes with a radius between 5 meters and 20 meters are most common for this subset. With higher resolution datasets, the minimum detectable size approaches observed sinkholes. However, to detect sinkholes, we need sufficient observable displacement. The next section will talk about the minimum detectable deformation.

4.2.3. Minimal Detectable Deformation and Detectability Power

The *minimum detectable size*, or MDS, shows the minimum size of a sinkhole we can detect at a particular location. Fig. 4.24 shows two maps with the minimum detectable deformation and the detectability power. More stable time series will provide a lower minimum detectable deformation. These maps provide a priori knowledge of how much deformation we need or what detectability power we can attribute to a specific deformation value. For example, when we detect anomalous behavior, how certain are we that the measured deformation can be attributed to an impending sinkhole?

4.2.4. Discussion

In this section, we first talked about implementing an artificial sinkhole into the dataset and its effect. We implemented a linear subsiding sinkhole and an accelerating subsiding sinkhole. When we look at the velocity maps of each implementation, we see a (slightly) decreasing linear trend. However, we can not make strong statements solely based on the velocity map without more definitive clues.

The second topic we researched is the minimum size we can detect. The size is dependent on the resolution of the observations. The observations are not homogeneously distributed, and thus, the detectable sizes highly vary. The histogram (Fig. 4.23) shows the distribution of the detectable sizes. Fig. 4.23 shows that 64% of the locations will have a minimum radius of 20 meters or smaller. 86% of the area has a minimum radius of 30 meters or lower. The MDS map indicates that at 91% of the locations on the map, an impending sinkhole can be detected with a minimum radius of 35 meters. The most common minimum detectable size is a radius of 10 meters.

The last topic showed how we implemented the minimum detectable deformation or detectability power. When we have detected anomalous behavior in an observation, we can combine it with this metric to put a probability on the detection.

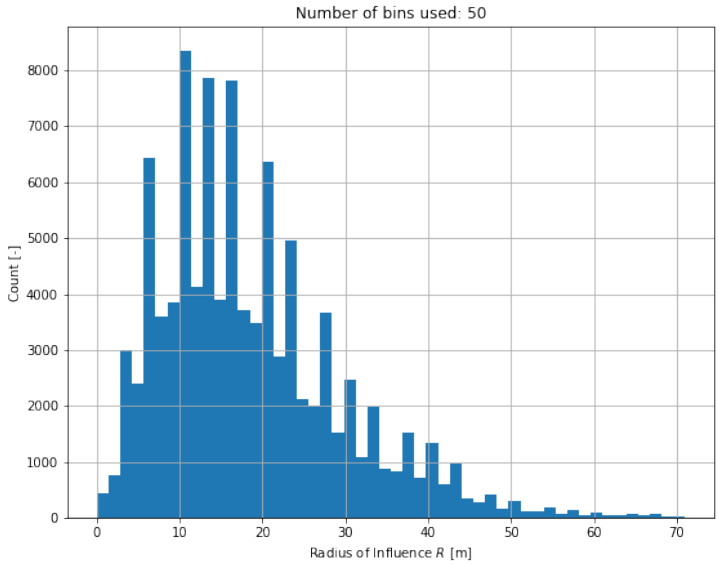


Figure 4.23: Histogram showing the distribution of minimum detectable sizes for potential impending sinkholes in the subset.

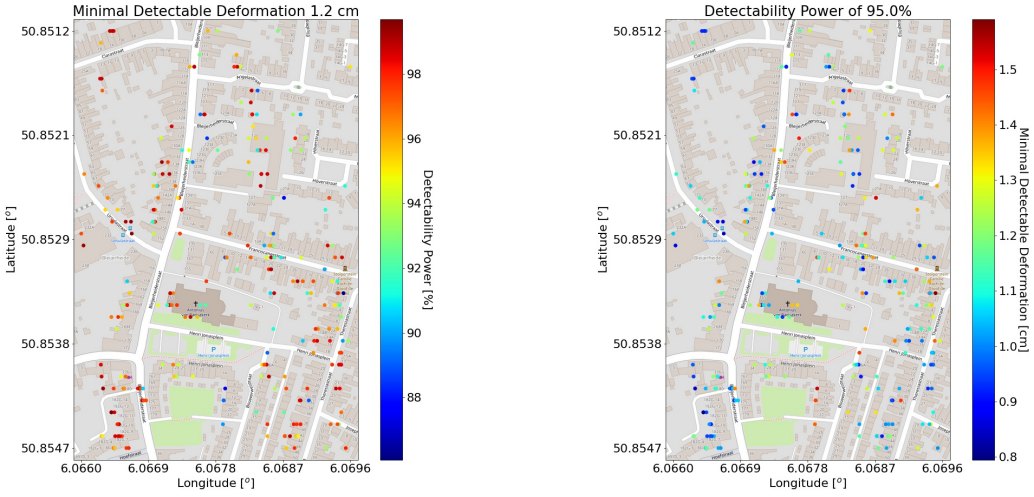


Figure 4.24: The left figure shows the minimum detectable deformation with a detectability power of 95% and a significance level of 5%. The right figure shows the detectability power with a minimum detectable deformation of 1.2 centimeter and a significance level of 5%.

4.3. Spatio-Temporal Strategy using Kinematic Model

This section will describe the results obtained from implementing the Spatio-temporal strategy performed with the kinematic model. In Section 3.3 we explained how we execute this strategy. We used the same subset as described in Section 4.2.1. In this subset, we know when the sinkhole signal is active, and thus we can anticipate when and where the sinkhole signal should be present. First, we quickly tested the strategy using arbitrary values for the number of epochs used for the kinematic model, the radius of influence, and the grid density. Then, we tested individually how each parameter influences the result. This part of the research aims to (1) investigate the feasibility of implementing the kinematic model and (2) identifying parameters influencing the results.

4.3.1. Implementation Spatio-temporal strategy

Here, we implemented the search algorithm. There are three parameters set:

1. The number of epochs used in this model is 10. Thus, the model computes the subsidence rate over ten epochs, which is about a range of 60 days depending on the epochs selected.

2. We set the radius of influence to 35 meters. Thirty-five meters is the radius used for the implemented sinkhole. We assumed that it is easier to set it to the same value initially.

3. We used a grid of 10 by 10, totaling 100 grid points. The distance between each posting (i.e., grid point) is 25 meters.

We used the filter functions for both the estimate ($\underline{\hat{x}}$) and the estimated variance ($Q_{\hat{x}}$). We implemented these filter options to reject unrealistic estimates that do not fall within our window of interest (e.g., we are not interested in ground heave). We have set the filter boundaries as follows:

1. We set the window for subsidence velocity estimates to 0 for the upper bound and -0.0006 for the lower bound. The upper bound is evident since we only are interested in subsidence behavior. We set a lower bound regarding our implemented sinkhole. We know that the implemented sinkhole has a maximum subsidence of 5 centimeters. Five centimeters in half a year amounts to roughly -0.0003 meters per day of maximum subsidence. Thus, we assume that the maximum measured subsidence does not exceed the -0.0003 meters per day. To be on the safe side, we multiplied the computed maximum subsidence by a factor of two, obtaining a lower bound of -0.0006 meters per day.

2. For the filter of the estimated variance, we were more liberal. We settled on a filter window between 0 and 0.0001 (1e-4). We obtained these values by trial-and-error, and these values seem to work best.

We visualize the results using two figures. The first (left) figure displays the marked observations and their corresponding index number. Each marked observation we color according to their estimated subsidence velocity. The second (right) figure shows a bar plot of each grid point and its subsidence velocity. We mark each bar with its index number. Fig. 4.25 shows the final result of the Spatio-temporal strategy in action. We see some marked observations, but observation 55 shows a lot of subsidence present in the area. The estimated subsidence is indeed from the implemented sinkhole. Therefore, using the Spatio-temporal strategy using the kinematic model can detect an impending sinkhole. However, we had to do much tweaking (especially with the filters) to get these results.

Now, we look at the three parameters mentioned earlier and look at their effect on the results. We took the date 08-10-2018 as the comparison date because it is one of the first dates to show the implemented sinkhole signal and also shows numerous false positives.

Influence of the number of epochs used

In this subsection, we take a closer look at using different epochs for our estimation. We assume that more epochs will provide less noise. Using only a few epochs, we assume it will give a lot of false positives.

Ten epochs are roughly sixty days. We know that the sinkhole is active for at least half a year. Therefore, we will first look at the results when we use twenty epochs. Twenty epochs correspond with roughly 120 days (~four months). Fig. 4.26 has the same setup as Fig. 4.25, but now we have

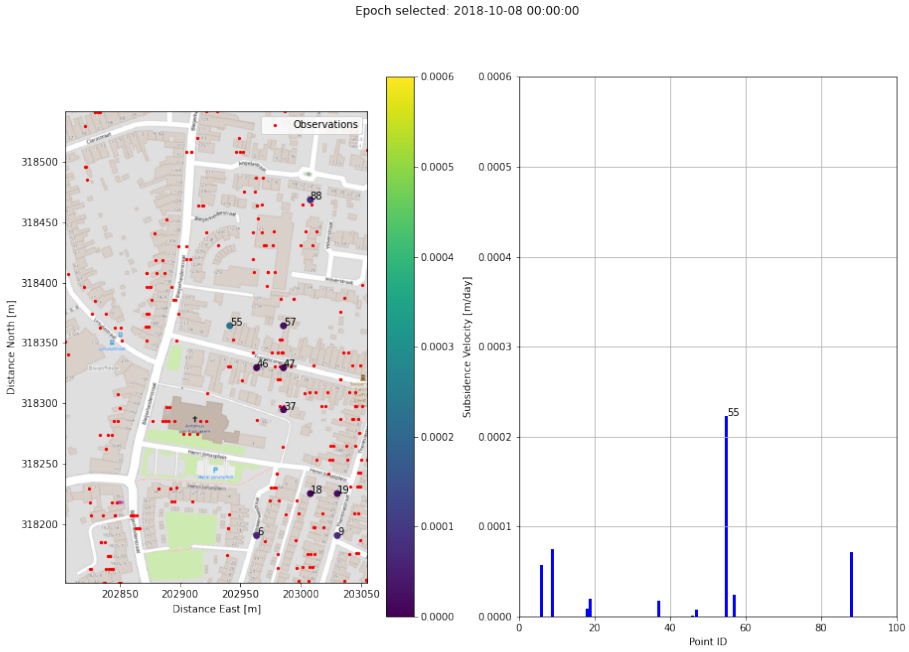


Figure 4.25: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 35 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

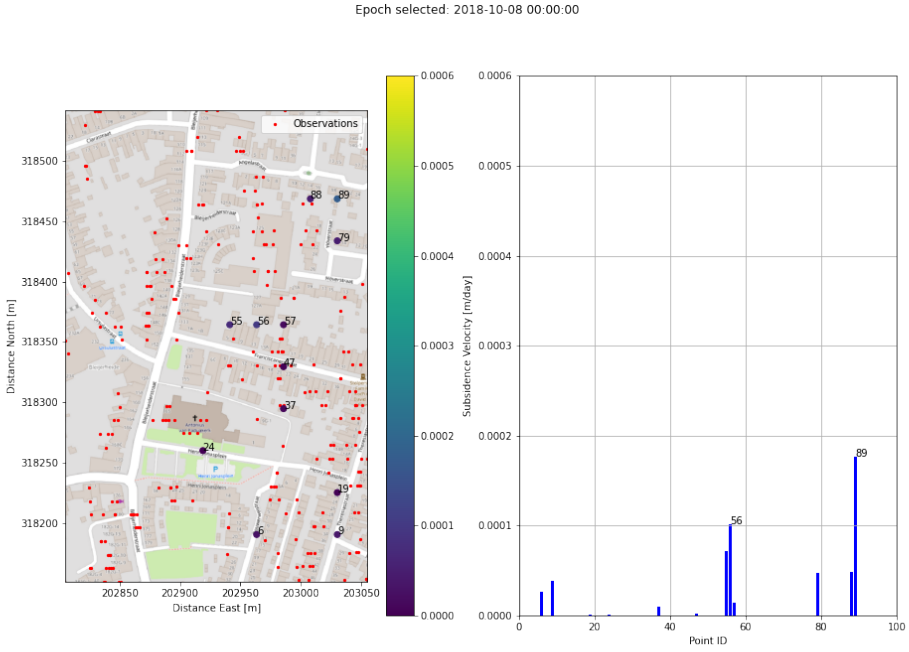


Figure 4.26: Results for the subsidence velocity estimation using 20 epochs, radius of influence of 35 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

used twenty epochs to estimate the subsidence velocity. We can see the same signal but with a smaller magnitude. Also, other marked observations popped up. A quick conclusion is that ten epochs better capture the signal of interest.

The next step is to use fewer days for the estimation. We try and use five epochs (about thirty days). The result we got was almost non-existing. Hence, we increase, only for this case, the upper bound of the filter for the estimated variance to 0.001 (1e-3). We increase the upper bound with a factor of ten. With the new limit, we obtain a fair amount of results. Fig. 4.27 shows the results for

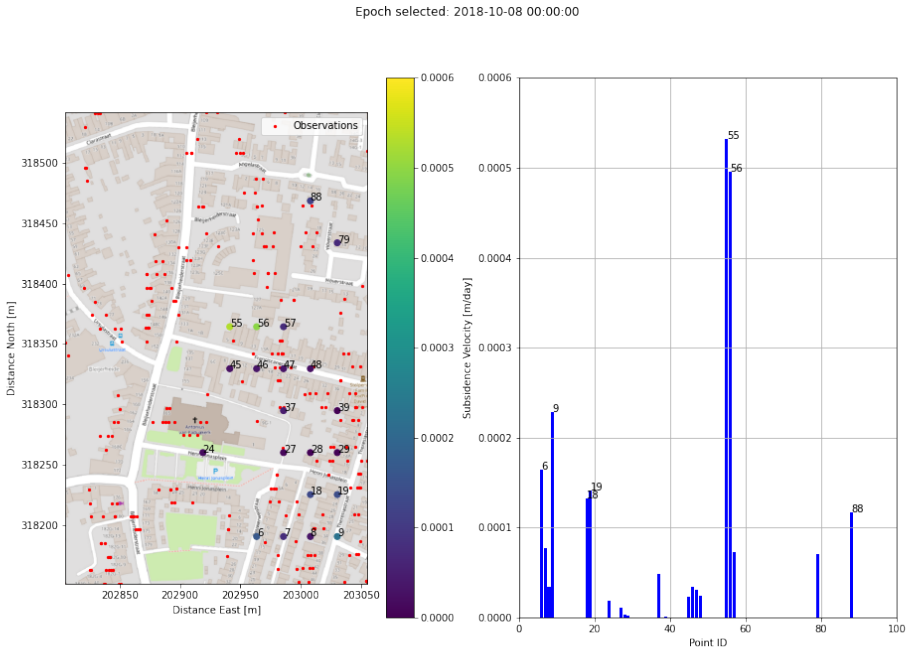


Figure 4.27: Results for the subsidence velocity estimation using 5 epochs, radius of influence of 35 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

using five epochs and an increase for the upper filter bound. The sinkhole signal is even more apparent in these figures but also increases the detection of other observations. The exciting part is that the affected points show a very high subsidence velocity. If we implement a threshold approach, we might be able to set up a detection algorithm.

Influence of the Radius of influence

In the next step, we look at the influence of selecting different radii. We pick five different radii for October 8. The radii that we tested are $R = 17$ m , $R = 25$ m, $R = 33$ m, $R = 40$ m, and $R = 60$ m. We start with the smallest radius of influence, $R = 17$ m. Fig. 4.28. By using smaller radii, we have to keep in mind the posting distance. When we investigate the area, we want the grid and the radius of influence to cover the whole area. We do not want uncovered areas which might host some points. In the case of $R = 17$ meters and a posting distance of 25 meters, we constantly have an area of around 1 m^2 that is not covered. The exposed area lies on the diagonal between the grid points.

On October 8, we do not see much and therefore conclude that we do not see any impending sinkhole in the area with a radius of influence of 17 meters. Increasing the radius to 25 meters gives us at least coverage of the whole area. Still, the radius is much smaller than the implemented sinkhole radius. In Fig. 4.29, we still do not see a lot of signals indicating an impending sinkhole. The results look similar to Fig. 4.28. We increase the radius again to almost the radius of the implemented sinkhole. We expect that we should see at least some indication of the impending sinkhole since the difference in radii is minimal. Fig. 4.30 shows the result, and to our disliking, we do not see anything out of the ordinary. We suspect that the grid size has to play a role here. We already looked at Fig. 4.25, and here we see a sinkhole signal at observation 55, which is the closest point to our implemented sinkhole.

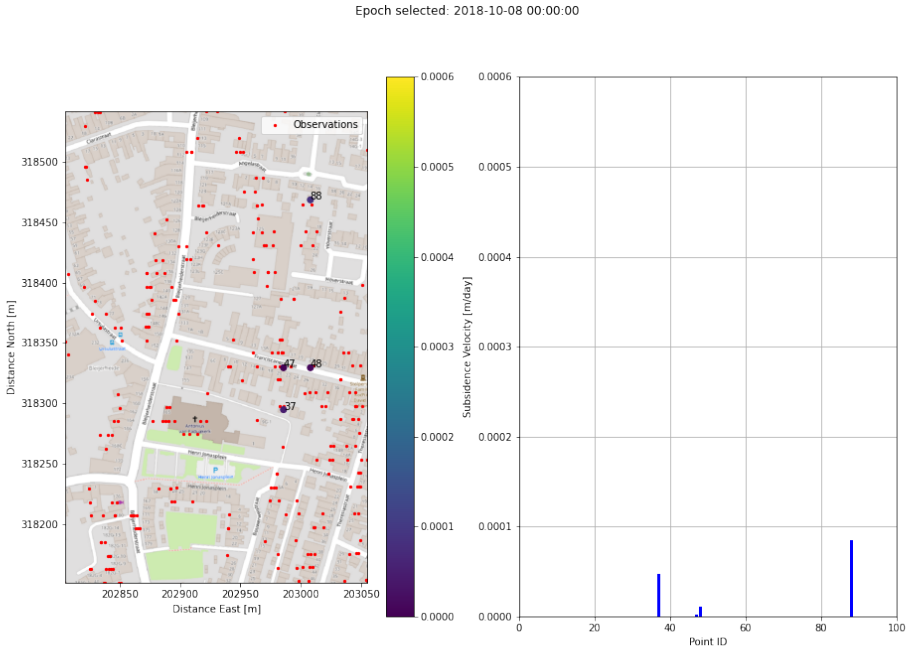


Figure 4.28: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 17 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

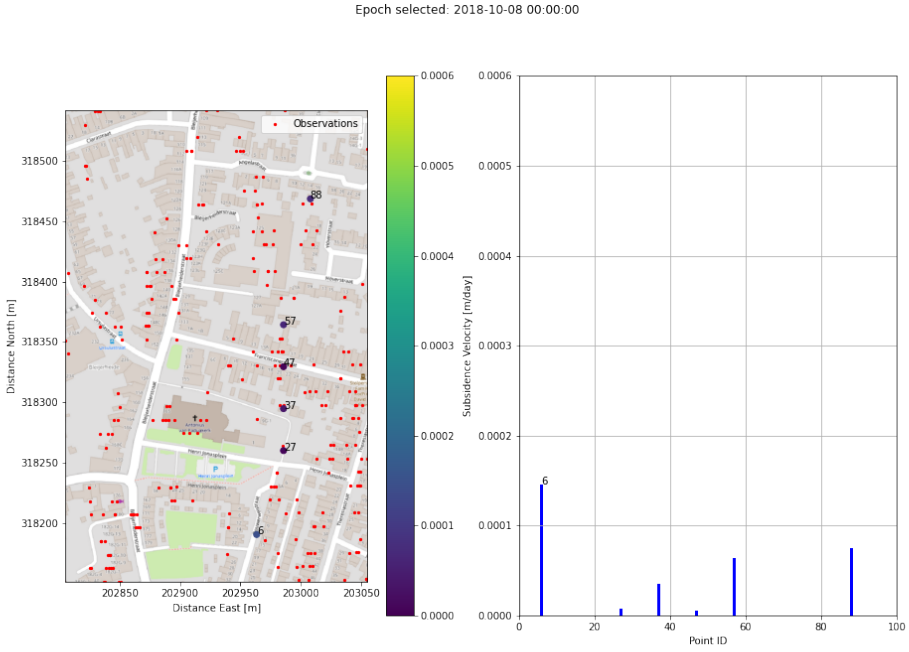


Figure 4.29: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 25 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

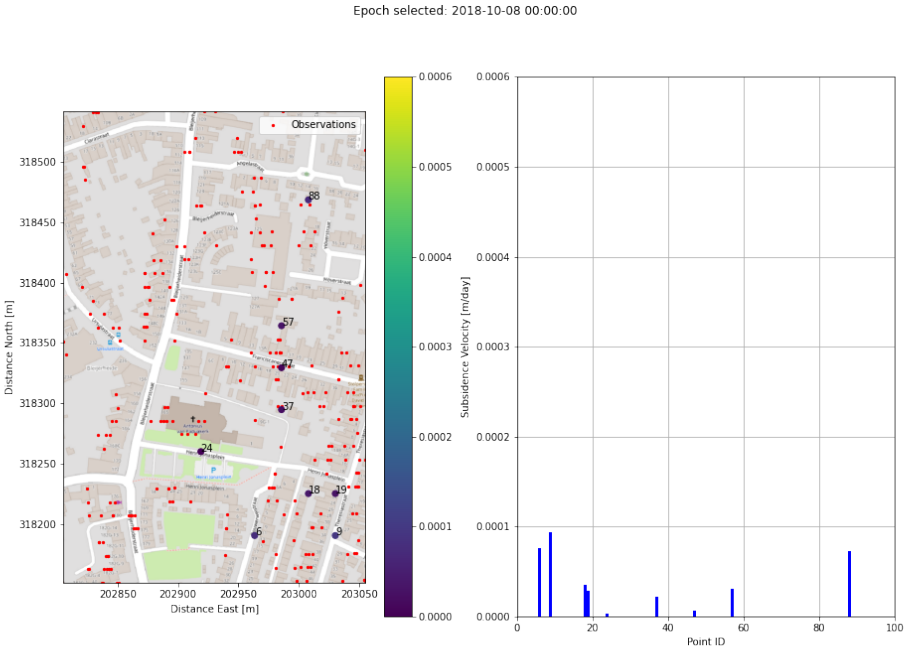


Figure 4.30: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 33 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

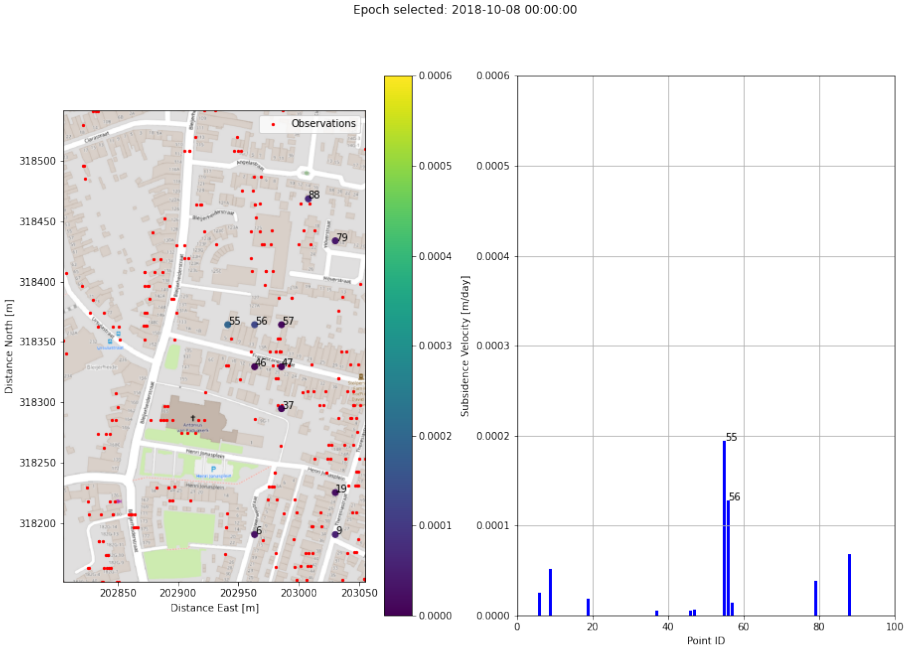


Figure 4.31: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 40 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

Fig. 4.31 shows the search for the impending sinkhole with a radius of influence of 40 meters. This figure clearly shows something happening at point 55 on the contrary to Fig. 4.30. We suspect that the posting is an issue here since the radius used in Fig. 4.31 is much larger than the radius used in Fig. 4.30. The last figure is Fig. 4.32. This radius of influence is almost twice as big as from the

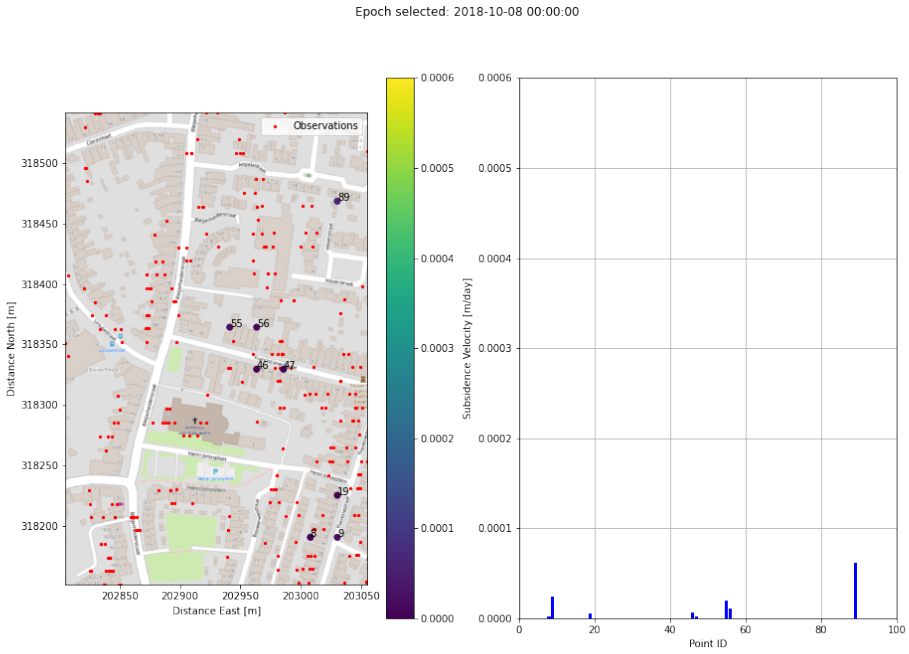


Figure 4.32: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 60 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

impending sinkhole. However, we still see that the algorithm picked up point 55, albeit not very strong.

Grid Density

The last parameter that we investigate is the grid density. Here, we use more points to evaluate the kinematic model and decrease the posting distance. The grid of 10-by-10 points had a posting every 25 meters (in x- and y-direction). We increase the density by defining a grid of 20-by-20, which has a posting of 12.5 meters. Finally, we obtain a grid of 40-by-40 with a posting distance of 6.3 meters. Here, we only looked at the effect the increased grid density has, and thus the number of epochs used and the radius of influence was set again to 10 and 35, respectively.

The increase of the grid density slows down the algorithm a lot. A grid of 10-by-10 only has 100 points to evaluate, while a grid of 20-by-20 had 400 points, and finally, 30-by-30 has 900 points to evaluate. Also, with the increase of grid points, the figure becomes messier since we show more data. Fig. 4.33 is the first grid density increase. As already mentioned, the posting distance decreased to 12.5 meters. We notice that a lot more points are marked, which is logical since we increase the number of points by a factor of four. With the increased grid density, we still observe the area of the impending sinkhole being marked, and by a lot more points. In the previous section (Section 4.3.1), we noticed that radii smaller than the implemented sinkhole were not detected while larger radii did pick up the impending sinkhole signal. Fig. 4.34 uses a smaller radius ($R = 25$ meters) and an increased grid density. Now, we do detect something at the location of the implemented sinkhole. We can thus conclude that the size of the grid can play an important role in detecting a particular size of a sinkhole. In the last figure, Fig. 4.35, we again increased the grid density to a total of 900 points. The posting distance decreases to 6.3 meters. Interestingly, with the increased grid density, we marked fewer points. However, the points are situated closer together. We removed the point numbering since this would clog up the figure and make it hard to interpret.

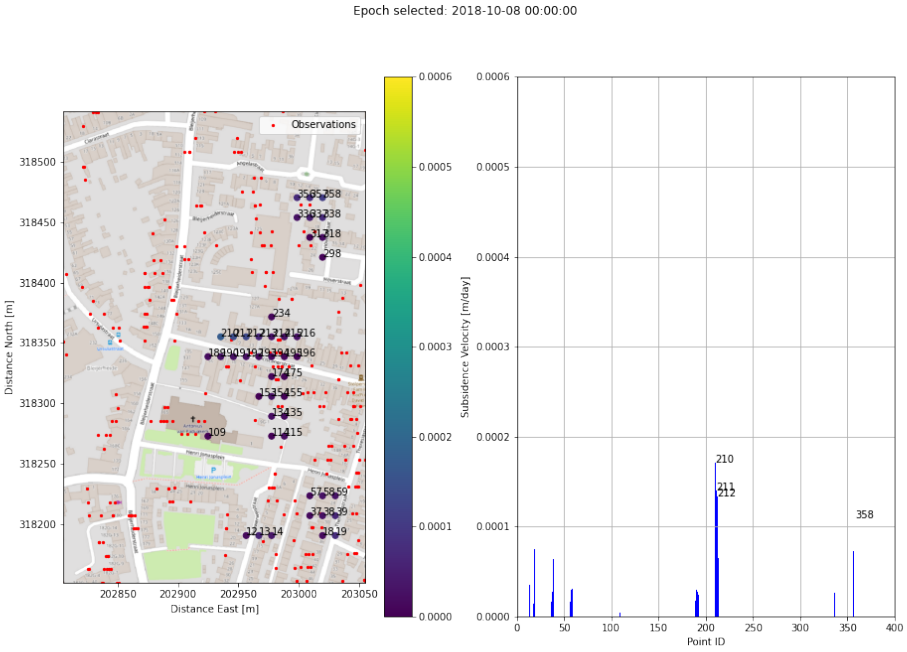


Figure 4.33: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 35 meters, and a grid density of 20 by 20 (posting every 12.5 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

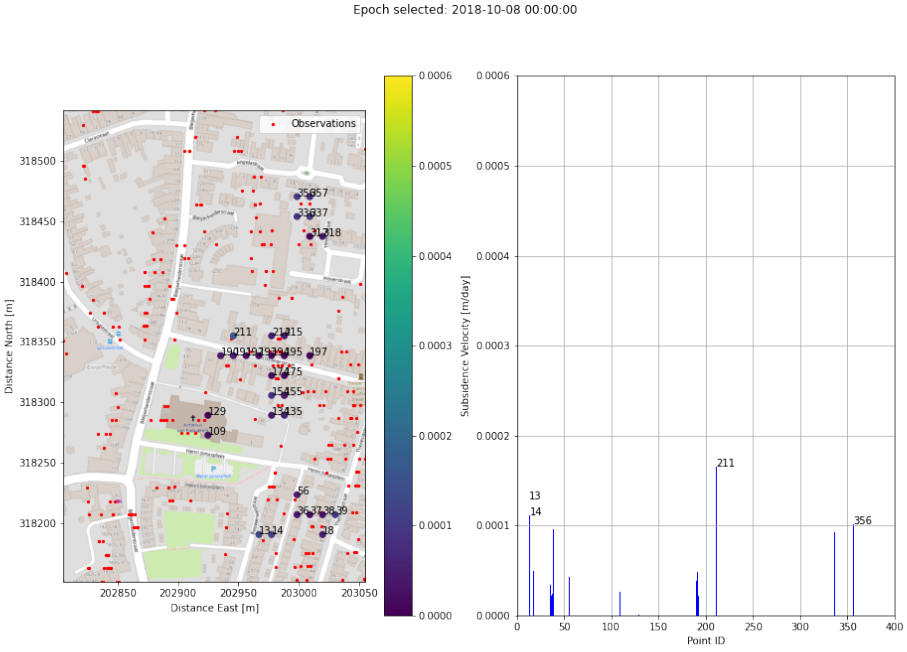


Figure 4.34: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 25 meters, and a grid density of 20 by 20 (posting every 12.5 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

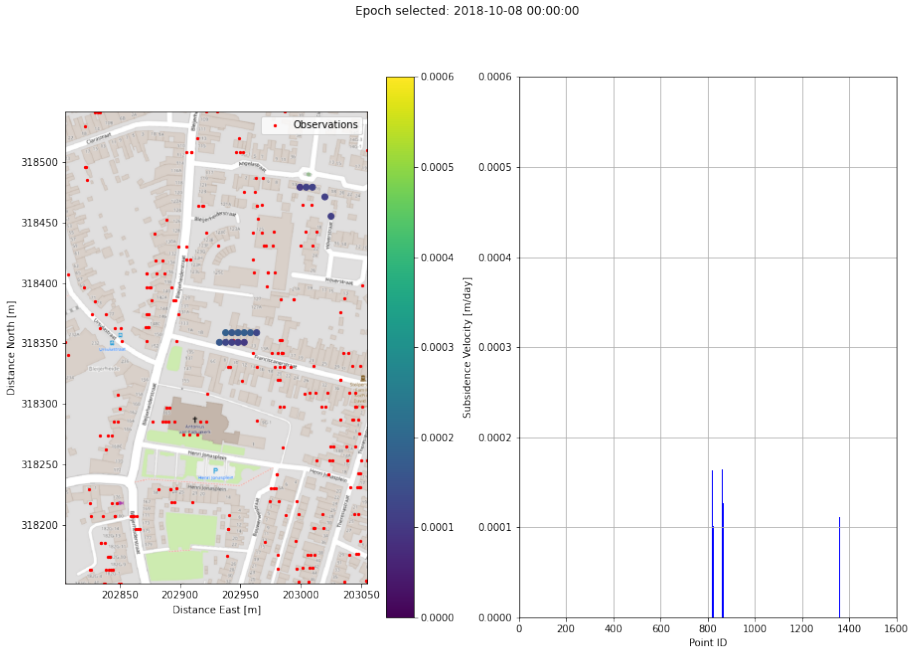


Figure 4.35: Results for the subsidence velocity estimation using 10 epochs, radius of influence of 35 meters, and a grid density of 40 by 40 (posting every 6.3 meters). The date is 8 October 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

4.3.2. Discussion

To get results using the spatio-temporal strategy requires a lot of rigorous tweaking and filtering of the estimates. Also, tweaking the number of epochs, radius of influence, and grid density significantly impacts the results. Still, the results are not always optimal, and sometimes the signal does not even get detected. All these factors make it very hard to predict an impending sinkhole accurately. Since we implemented an artificial sinkhole in the dataset, we know where to look. Therefore, it sometimes feels more of a game of filtering out false positives. If we do not know how the sinkhole signal would form, it would be tough to tweak this algorithm to detect an impending sinkhole.

We know that the subsidence velocity is 0.0003 meters a day, but the estimates seem to bounce around this value. Most of the time, the estimate is lower. We conclude that the estimated velocity is unreliable, and we can at most predict a probable velocity range. Changing the number of epochs does not lose the impending sinkhole signal. It mainly affects the magnitude and the number of false positives found. Using five epochs seems to be interesting because most sinkholes occur within a short time window. However, obtaining the results was only possible by adjusting a filter. What we observe is interesting. We detect a stronger signal for the affected points, but sometimes the signal completely disappears. It looks like a good strategy, but when looking at Fig. 4.36, we see that it can also produce a result containing much noise. However, if one would develop a way to identify or reduce these false positives, then this set of parameters might harbor potential.

We detect the sinkhole using multiple R values. It means that the exact radius is hard to determine, and we can only obtain a range. A strange discovery is that a radius of 33 meters does not detect the sinkhole signal, while a radius of 40 meters does. The current explanation is that the grid density is inadequate to use smaller radii, and it seems to be supported by Fig. 4.34.

For computational reasons, we want to make the grid as sparse as possible. However, the drawback is that the detection of sinkholes tends to favor larger sizes. It might not be a problem if we only want to detect impending sinkholes, but it might impede obtaining more realistic results. The grid density also affects the minimal detectable size. We need at least a few observations close to a grid point before computing a reliable estimate. Increasing the posting distance together with a decreasing radius of influence might inadvertently exclude many observations.

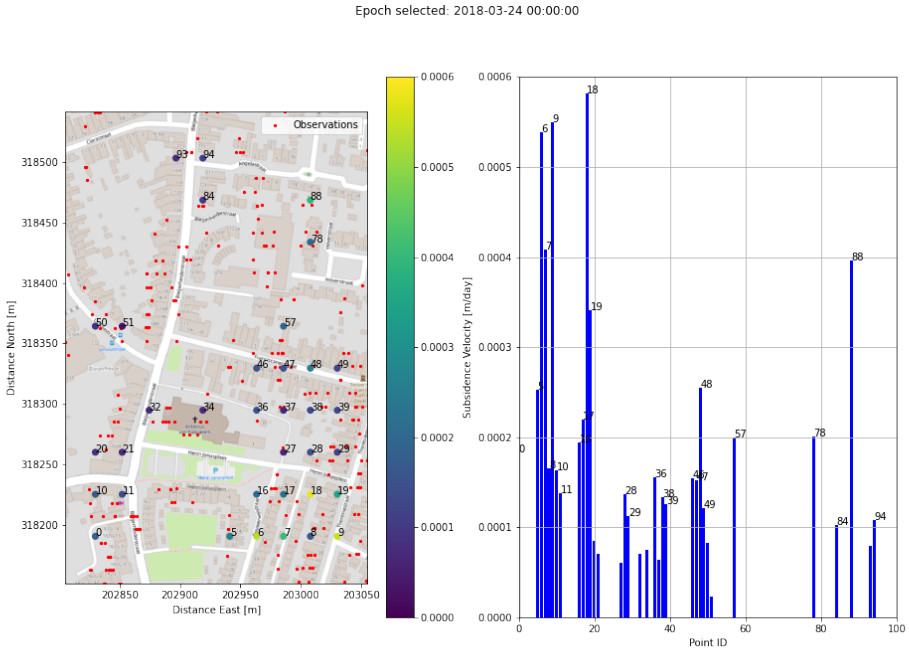


Figure 4.36: Results for the subsidence velocity estimation using 5 epochs, radius of influence of 35 meters, and a grid density of 10 by 10 (posting every 25 meters). The date is 24 March 2018. The left image is a top view with the estimates not filtered out. The right image is are the bar plots of the estimates.

False Positives

We know when and where the sinkhole signal should pop up. Marked observations outside this time window, are labeled false positives. It makes our search more manageable. We endure many false positives, and luckily most of them have a minimal subsidence velocity. Sometimes this is not the case, and it forces us to find new ways to identify these false positives. We notice that a group of points in the top right corner is marked relatively more often. However, points too close to the edges might indicate false positives since shifting the subset would reveal more observations to correct the estimates.

Another theory is that false positives are inherent to the dataset. When using multiple sensors or multiple tracks, these occasional marks can get filtered out.

4.4. Temporal strategy using Arcs

In Section 3.4 we describe the arc-based methodology to detect anomalous behaving points systematically. This section applies the methodology to a subset whereby we superimposed an artificial impending sinkhole pattern on the data. We based the impending sinkhole pattern on the kinematic model (Section 3.1). The methodology consists of two steps. The first step is the characterization of an arc. Here, we use a linear model and sufficient epochs to characterize an arc. We also make a statement about how we think the model should be updated. The second step shows the detecting procedure of the impending sinkhole. We visualize this performance using a top-view plot and a bar graph. The idea is that an anomalous behaving point has more arcs with a high flag count connected, which would be visible in the bar graph of the total flag count. This part of the research aims to (i) develop an alternative methodology for finding sinkholes and (ii) identifying some key parameters when applying this methodology.

4.4.1. Arc Characterization

An essential aspect of the arc-based detection method is the characterization of the computed arcs. Ideally, we want to characterize the arc so that only anomalous observations are flagged. Unfortunately, the characterization of an arc tends to be more difficult in practice due to errors such as unwrapping errors, outliers, (seasonal) trends, point quality, physical change (e.g., placing of a dormer), weather,

or other sources provoking volatile behavior.

However, it seems that a linear model with a sufficient confidence interval encompasses most of the arc behavior. Albeit, if we chose the model size sufficiently and the number of to be flagged observations concisely. Fig. 4.37 shows three figures. We have based the selection of the three

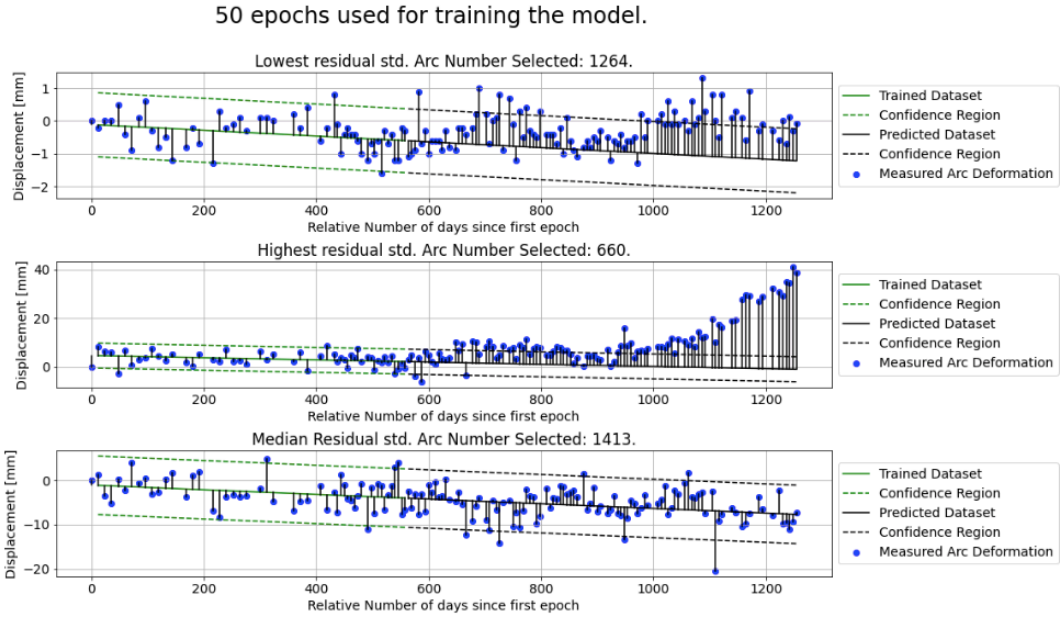


Figure 4.37: Three plots showing arc time series where we used 50 epochs to train our model. The top plot shows the time series with the lowest standard deviation of the residuals. The middle plot shows the highest standard deviation, and the bottom plot shows the median standard deviation.

figures on the residuals. We used all the observations in the computation for the residuals, and from the residuals, we computed the standard deviation. We then sort every arc to the computed standard deviation. The first (top) figure is the displacement time series with the lowest standard deviation. The second (middle) figure shows the highest standard deviation. As is visible in the figure, the highest standard deviation includes the point most affected by the implemented impending sinkhole. The third (bottom) figure shows the time series with the middle (median) standard deviation in the sorted series. Fig. 4.38 shows two histograms containing statistics about the arcs. The left figure shows a histogram

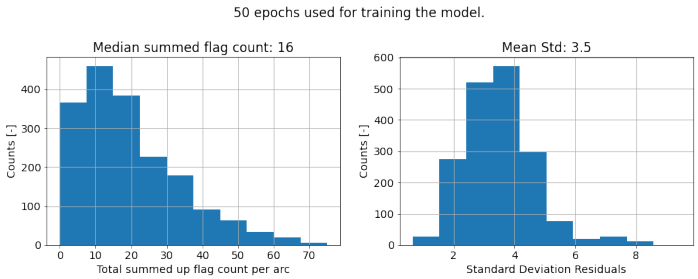


Figure 4.38: Two histograms showing the distribution of the total flag count of an arc, and the distribution of the standard deviation per model.

of the summed-up flag count per arc (without short-term memory retention, see Section 4.4.2). The summed-up flag count is the total number of times we have flagged the arc in the time series. We see that an average arc is flagged 16 times. We chose to use a median instead of a mean because we expect the outliers to be relatively large. The right figure shows a histogram of the standard deviations of the arcs. The average standard deviation is 3.5.

Having 50 epochs to apply the model on is already a large time span. In this case, 50 epochs stand for roughly 552 days, thus around 1.5 years. Still, we would expect that increasing the number of epochs will decrease the general arc flags and the standard deviation. Fig. 4.39 shows the same

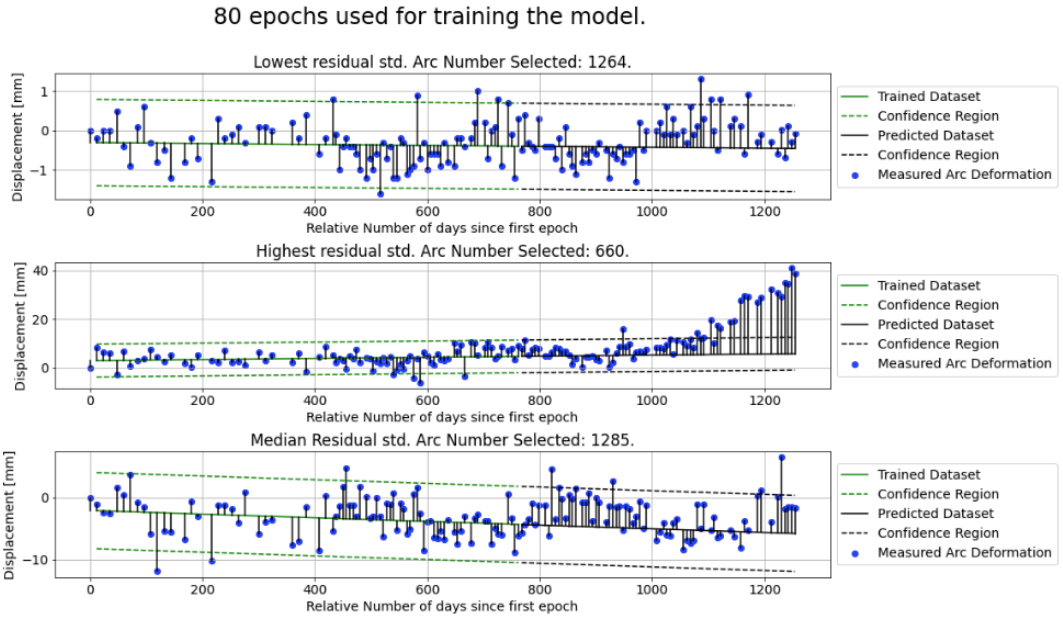


Figure 4.39: Three plots showing arc time series where we used 80 epochs to train our model. The top plot shows the time series with the lowest standard deviation of the residuals. The middle plot shows the highest standard deviation, and the bottom plot shows the median standard deviation.

three different time series but now the model used 80 epochs. Fig. 4.40 shows the histogram with

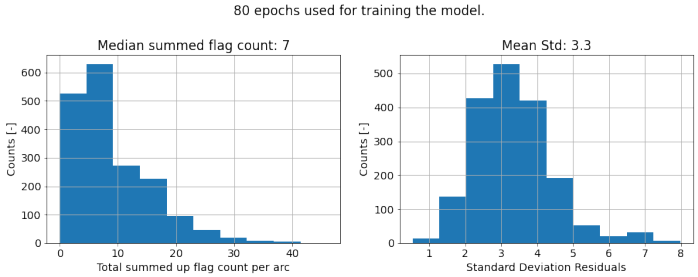


Figure 4.40: Two histograms showing the distribution of the total flag count of an arc, and the distribution of the standard deviation per model.

the summed-up flagged arcs and the standard deviation. Again, we see a clear improvement in both histograms meaning that more epochs indeed improves the model. Thus, using fewer epochs will produce worse results. Fig. 4.41 shows the results when using only 30 epochs to fit our model. The results are worse than in Fig. 4.37 and Fig. 4.39. In the histogram of Fig. 4.42, it is even better visible that the results of using fewer epochs produce a worse fit.

Therefore, a quick conclusion is that the quality of the model depends on the number of used epochs. However, characterizing the arc might also be influenced by the number of days between the last epoch in the model and the newly incoming observation. We investigate the relationship between the number of flagged observations and the model size used in the following steps. The two metrics to investigate the performance of the model are the median summed-up flag count and the standard deviation of the residuals. We compute the residuals using the whole time series. In Fig. 4.43 we automated the process of testing the model size. We kept the number of flagged observations the same per figure while changing the model size to incorporate more epochs. Using the metrics median summed-up flag count and the standard deviation, we can draw a few conclusions on the preferable model size and the preferable number of flagged observations before updating the model.

From experience, we flag arcs continuously and seemingly randomly, and most of the time only once. We can also recognize that in Fig. 4.43 since it converges to a specific value and not zero.

We notice that fewer observations flagged will provide a lower median flag count. We would expect

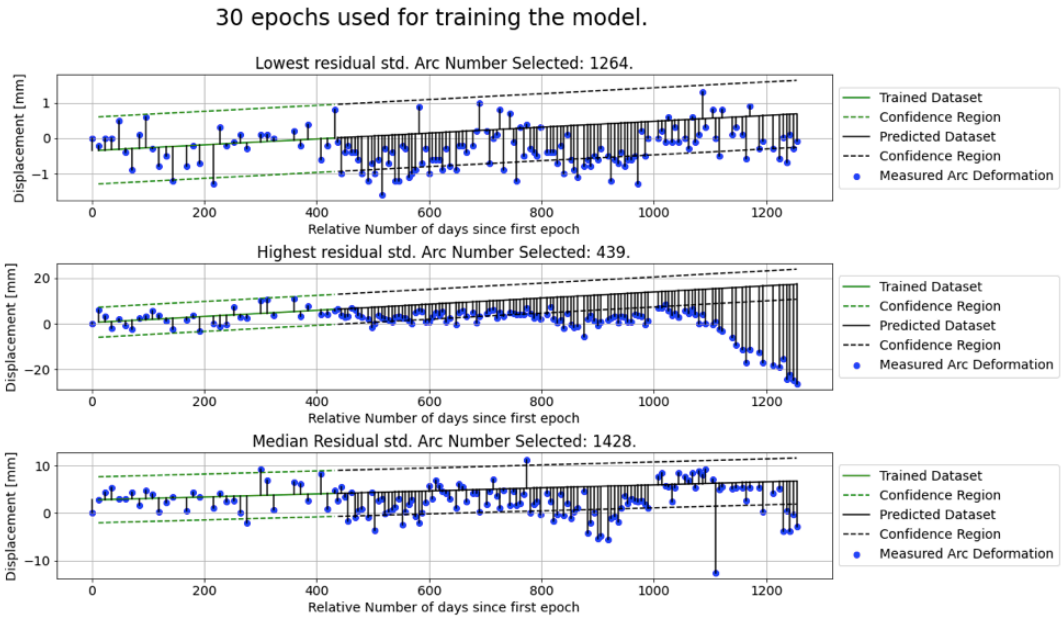


Figure 4.41: Three plots showing arc time series where we used 80 epochs to train our model. The top plot shows the time series with the lowest standard deviation of the residuals. The middle plot shows the highest standard deviation, and the bottom plot shows the median standard deviation.

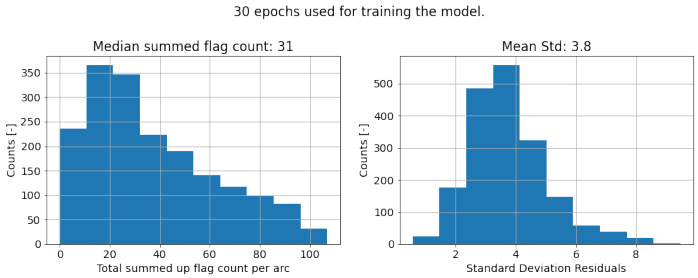


Figure 4.42: Two histograms showing the distribution of the total flag count of an arc, and the distribution of the standard deviation per model.

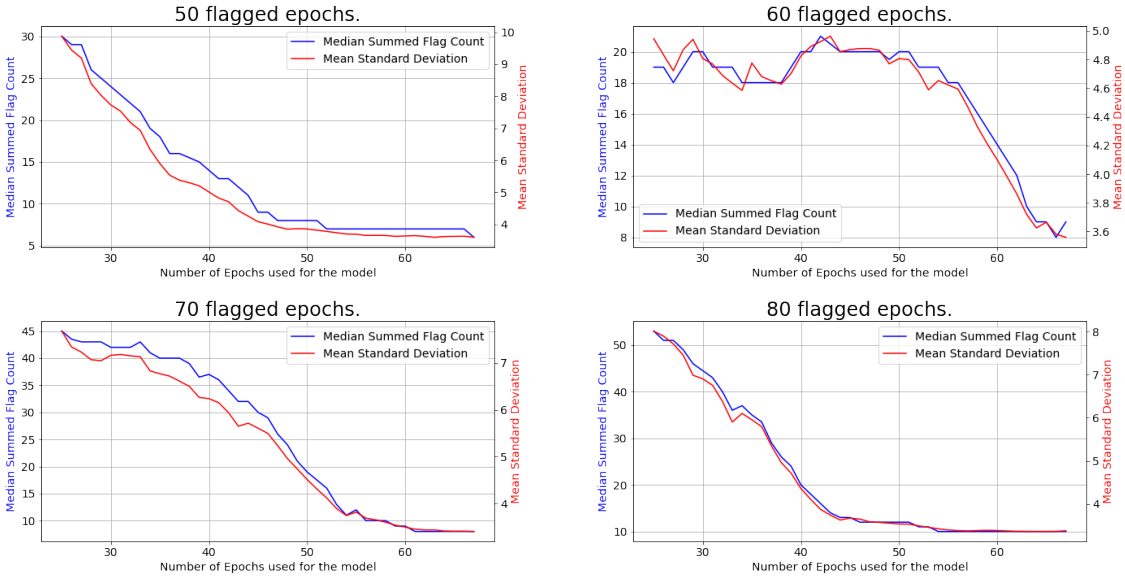


Figure 4.43: Figures showing the number of to be flagged observations relative to the model size. We chose a model size and then look at the median summed flag count and the std of the number of the to be flagged epochs. We do this for 50, 60, 70, and 80 to be flagged epochs.

this since observations are closer to the model. Also, a minimum of 60 epochs for the model is preferred. Sixty epochs are roughly one and a half years of displacement time series data. Further, one year of displacement data also covers the seasonal trend.

Discussion Update Strategies

The next step is to devise a strategy that implements new observations into the characterization model. We assumed that nothing is happening in our subset and that the only sinkhole signal is the implemented artificial sinkhole signal.

To get a decent characterization, we need at least 50 epochs (1.5 years) and preferably a little more (i.e., $60+$ epochs). This conclusion was drawn by looking at Fig. 4.43. By trial and error, we found that using more or fewer epochs does not significantly impact the computation time, fortunately, and therefore does not pose a big issue.

Another aspect is to change from a linear model to a seasonal or higher polynomial model. A (slight) seasonal trend is visible in some of the time series. We expect the arc creation process to filter out most of this signal because a seasonal trend in two observations would get subtracted from the newly obtained time series. However, in some cases, a seasonal model might still better characterize the arc.

A simple model updating scheme is to add an observation to the model that we did not flag. Thus, the model will get extended with *approved* observations. In this way, the incoming observation also lays close to the model. Nevertheless, a deviating trend might be incorporated into the model too soon. An option can be to lag behind several epochs before the approved observation gets appended to the model. We can think of many other updating schemes. To test the effectiveness of the update scheme, the median flag count might be a good metric to keep in mind. The most important aspect of an updating scheme is to use sufficient (i.e., 1+ years of epochs) to characterize the displacement time series. The temporal size of the model seems to have the most influence on the characterization process.

4.4.2. Arc-Based Detection Method Implementation

This subsection shows how an arc-based detection method would look like and how anomalous points can be detected.

In the previous section, we concluded that ideally, we use at least 60 epochs for the model for characterizing an arc and use around 50 epochs to flag. During the flagging procedure, we introduce a short-term memory in the shape of a flag count. The flag count tracks the number of times the flagging procedure flags an arc in succession. We use the flag count also for coloring the arc in the

visualization. We set a maximum value for the flag count to four. The first time we flag an arc, we do not show it, and we do not use it for the coloring scheme. The reason is that with so many arcs, single outliers tend to be appearing too often, which only clutters the image. Later, we use the flag count to create the total flagged value per observation. The total flagged value represents all the flagged arcs going towards an observation, and we visualize this using a bar plot.

We currently visualize the flagging procedure using a short clip; however, this is not possible to show within this thesis. Therefore, we show snapshots from the clip. The actual link to the clip can be found in the Appendix A.

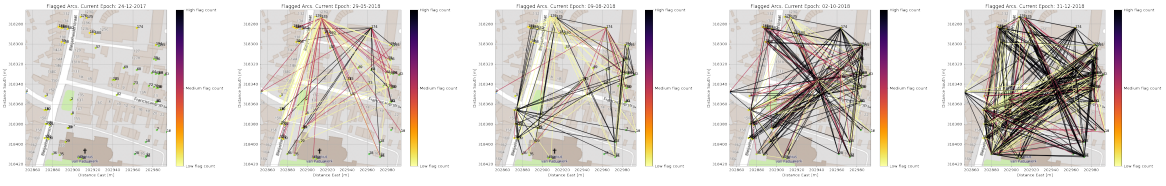


Figure 4.44: Top view of the time series consisting of 5 epochs. At each epoch we show the flagged arcs colored corresponding to their flag. The first epoch is 24-12-2017, the second epoch is 29-05-2018, the third epoch is 09-08-2018, the fourth epoch is 02-10-2018, and the last epoch is 31-12-2018.

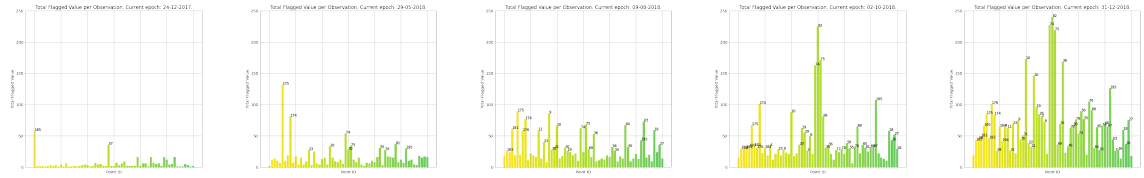


Figure 4.45: Bar plot of the time series consisting of 5 epochs. At each epoch we show the total flagged value per observation. We start numbering the bars above a value of 20. The epochs are consistent with Fig. 4.44.

From Fig. 4.44 and Fig. 4.45, we see that more and more arcs are getting flagged towards the end of the time series. Fig. 4.44 tends to get very messy at this point, while in Fig. 4.45 we see some observations hinting at anomalous behavior.

Since we implemented an artificial sinkhole, we know which observations are most affected, making it a little easier to spot anomalous behavior. For example, we know that observations 82, 74, and 75 are most affected, and see it stick out in Fig. 4.45.

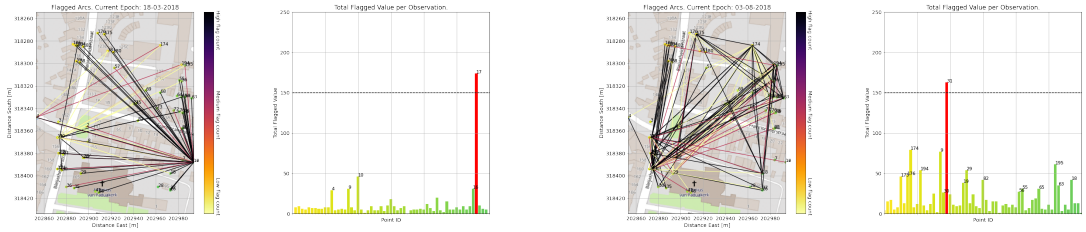


Figure 4.46: Two epochs showing two random observations above the set threshold indicating anomalous behavior.

Threshold for Minimum Total Flagged Value

To detect these anomalous points, we would like to set a threshold. This threshold suggests that above a set number, the procedure will mark the observation as behaving anomalously.

Let us say we put the detection threshold at a total flagged value of 150. We define this value by looking at the observations' regional behavior, especially the bar plot. We expect to define a new threshold for every new region and every new sensor because each dataset or sensor might be affected

by other types of errors or noise. However, the short-term memory has biggest effect on the threshold. In this subset, there are two moments when two random observations are above the threshold.

Fig. 4.46 shows the two epochs when the detection threshold would classify these observations as behaving anomalously (i.e., above a threshold value of 150). However, the observations are only classified for one epoch, while the affected observations (i.e., observations 82, 74, and 75) are consistently flagged. The table appendix B shows the marked points (points exceeding the 150 TFV threshold) in

Table 4.6: Shortend table of the marked Total Flagged Values obtained from running the arc-based strategy. The full table can be found in Appendix B.

Epoch Date	ID	Total Flagged Value
2018-03-18	17	174
2018-08-03	31	163
2018-09-14	82	179
2018-09-26	82	202
2018-10-02	74	163
2018-10-02	82	225
2018-10-02	75	171

the whole subset. There are two randomly marked point, but are marked only once. At 14 September we notice that point 82 is getting marked and stays marked. On 2 October the two neighboring points are marked also denoting that something anomalous is happening at this particular location.

We can find the details of the implemented sinkhole signal in Table 3.1 On 14 September 2018, observation number 82 is flagged, and on 2 October 2018, the procedure flags the three most affected observations.

Table 4.7: Displacements introduced by the artificial sinkhole. This is only the sinkhole signal and not the time series displacement.

Point ID	Distance [m]	Displacement [mm]			Detectability Power [%]			Total Flagged Value [-]		
		14-Sept	26-Sept	02-Oct	14-Sept	26-Sept	02-Oct	14-Sept	26-Sept	02-Oct
82	15	-12	-14	-15	90	100	100	179	202	225
75	20	-7	-8	-9	49	97	100	99	140	171
74	19	-8	-9	-9	53	96	100	77	121	163
2	33	-1	-2	-2	13	5	21	13	14	26
73	31	-2	-2	-2	8	14	21	11	9	21

Table 4.7 shows the various observations that the sinkhole signal influences. The *Displacement* column shows the implemented displacement by the subsidence trough. We compute the Detectability Power (DP) by using the linear trend computed in Section 4.4.1. The difference the incoming observation has with the trend is used as the Minimum Detectable Deformation. The significance level was set to 5% (used throughout the thesis) and using the methodology explained in Section 3.2.4 we compute the DP. The table ends with the Total Flagged Value we gave each point.

4.4.3. Discussion

We introduced a flagging procedure with a short-term memory in the shape of a flag count. The benefits of using the flag count is that consistently anomalous behaving arcs are magnified. The flag count is then used to compute the Total Flagged Value (TFV) which is in essence the summation of all the flagged arcs towards a single measurement point.

The TFV has showed that using a threshold value is possible to judge whether a point is behaving anomalous. Here, we've taken a threshold value of 150 to mark points anomalous. This value was obtained by looking at the regional behavior. The threshold is highly influence by the short-term memory. Longer memories will need a higher threshold since the TFV will be substantially larger. Using the discussed tools, we detected two anomalous points (point ID 17 and 31), but these were only marked once and thus assumed to be false positives. Starting 14 September, point 82 is marked and stays marked. This is roughly 3 months after the impending sinkhole signal started. Point 82 lies roughly 15 meters from the center of the subsidence trough is already detected with 12 mm of

implemented displacement. On 2 October point 75 and 74 are also marked (see Table. 4.7). Table. 4.7 also shows the corresponding Detectability Power (DP) with the significance level set to 5%. Both the TFV and DP show a similar trend. Flagged points also exhibit a very high DP. This strengthens the detection of an impending sinkhole. The TFV is preferred over the DP, since the TFV uses the information of all the produced arcs and therefore the information in all the points within the used subset.

5

Conclusion and Recommendations

This study assessed the potential of using PSI time series to detect impending sinkholes. The key findings this research has produced are discussed in Section 5.1. The key findings are split up into the sub-questions that will subsequently be answered. During the course of this thesis, the arc-based method was developed creating the last sub-question in the research. After the sub-questions, a general conclusion for this research can be found in Section 5.2. Finally, this research also produced three recommendations for future research and implementation of the developed methods. The recommendation can be found in Chapter 6.

5.1. Key Findings

How does a subsurface cavity manifests itself onto the surface, and how can we model the manifestation?

A subsurface cavity causes a subsidence trough at the surface. The generated trough was modeled using the kinematic model. The kinematic model generated various insights such as ...

Another insight into the manifestation of a subsurface cavity came from approaching the generated subsidence from two different viewpoints. The Eulerian viewpoint only concerns itself with vertical displacement while the Lagrangian viewpoint follows the path of a particle. The two viewpoints are compared using geometric model (i.e., kinematic model) and a physical model (i.e., Mogi point source model).

From the Lagrangian viewpoint it is expected that a rising subsurface cavity would cause accelerated subsidence since each discrete point would move towards the center making the points more subjective to the cavity's forces. The effect of accelerated subsidence was not visible when iteratively small volume changes would occur. An accelerating behavior was only visible when the changed volume increased to unrealistic sizes. The effect of moving ground particles will therefore not have a large impact on the subsidence rate and thus the kinematic model is a good approximation of the subsidence trough. An interesting discovery in this thesis was that the dominant direction of the displacement changes at roughly $\frac{2}{3}$ of the radius of influence. This holds for both the kinematic and the Mogi model. It means that depression troughs would exhibit more horizontal displacement at the edges.

The last insight came from the comparison of the kinematic model and the Mogi model to some real-world cases. The fit of the models start bad but get better as the trough deepens. In the Dead Sea case accelerated subsidence was observed and showed that the last frame (just before catastrophic collapse) showed worse results. The bad fit in the last frame is suspected to be caused by small local failures since full on failure would already happen shortly after.

Over the lifetime of an evolving subsidence trough the kinematic model showed to be adequate at modeling an impending sinkhole.

How do the spatial and temporal time series characteristics influence impending sinkhole detection?

The spatial characteristics of the time series is illustrated using a map on where the Minimum Detectable Size (MDS) is visible. The insight this map provides are the areas where there is a higher chance of detecting impending sinkholes due to the ability of detecting smaller sized impending sinkholes. It provides a priori knowledge as to the effectiveness of the spatio-temporal strategy (Section 4.3.1) at any specific location. The histogram (Fig. 4.23) of the MDS map shows that at 91% of the locations on the map, an impending sinkhole can be detected with a minimum radius of 35 meters. This size corresponds to the size of the implemented impending sinkhole.

The temporal characteristics can be typified with the Minimum Detectable Deformation or Detectability Power. Using the methodology described in Section 3.2.4 the insight into how much deformation can be detected given a specific detectability power and significance level was developed. It adds to the ability to provide a probability when anomalous behavior is detected in a specific observation.

How can a spatio-temporal strategy based on a kinematic model improve early detection of an impending sinkhole?

With the spatio-temporal strategy, we implemented a grid-search approach that tries to fit the kinematic model on each posting. The results were plentiful and therefore it was hard to distinguish the actual signal from false signals. Precursory subsidence usually has a short time span and therefore it is of most interest to use as few epochs as possible. However, fewer epochs will overlook the actual signal sometimes, we need to adjust the filter, and produces more false positives. Another important factor is the grid density. The denser the grid, the better the results will be.

We noticed persistent groups of marked grid points which highlights the location of the impending sinkhole. However, this methodology is prone to false positive groups thus for a definitive warning other clues must be gathered. The spatio-temporal strategy provides a range of possible surface expression sizes in the form of radii of influences. Estimating a range of possible sinkhole expression sizes is very useful for assessing the severity of the impending sinkhole and contributes to the improvement of the early detection. The spatio-temporal strategy also provides a general location of the impending sinkhole since the results will be clustered around a specific location.

How can an arc-based temporal strategy improve the early detection of an impending sinkhole?

The arc-based approach consists of two parts. The first part is the characterization of an arc. We need to characterize the individual arcs to detect anomalous behavior in these arcs. We characterized the arc with a linear model consisting of at least 50 epochs but preferably 60 epochs or more for Sentinel data. The insight we obtained from the results is that the model size comes down to more than 1.5 years of time series estimates for the characterization process using Sentinel-1 data. With smaller model sizes, false detections will increase. We tested the ideal model size using the defined Median Flag Count (MFC). This metric indicates the median total number of flags per arc. The lower the value, the better the model predicts the arc' s behavior. The insight we obtained is that the MFC commonly plateaus out at a certain value. The number of epochs needed for the plateau value is the optimum number of epochs needed for the model. New strategies to update the model can use the MFC to evaluate the performance of the newly developed strategy.

The second part consist of the actual detection of anomalous behavior. The flag counts of all arcs for a specific measurement point are summed up and to create the Total Flagged Value (TFV) per measurement point. We can then implement a threshold that detects when the TFV exceeds the threshold and thus exhibits anomalous behavior.

The arc-based temporal strategy can improve the early detection of impending sinkholes. Section 3.4.3 and 4.4.2 shows how we proposed a new method, based on arcs, to utilize all the information in the subset to detect anomalous behaving points.

We detected anomalous behavior by setting a threshold of 150 for the TFV. Two points were falsely flagged, but quickly discarded. Roughly 3 months after the subsidence trough started developing, anomalous behavior was detected in point 82. The detection started on 14 September 2018 and lasted

to the end of the year. On 2 October points 74 and 75 were also marked therefore strengthening the case of detection. Table 4.7 also showed the Detectability Power (DP). The DP concurs with the TFV meaning that marked points also show a very high DP. The DP looks similar to the TFV, however we prefer to use the TFV since it uses information from all the other points.

5.2. General Conclusion

In general, we conclude that the kinematic model is adequate to model the manifestation of a subsurface cavity onto the surface. We came to this conclusion by comparing two different viewpoints and compare the kinematic model to real-world cases. We found that inspecting a dataset's spatial aspects can be achieved by creating the Minimum Detectable Size maps. The temporal elements are characterized by computing the Minimum Detectable Displacement given the Detectability Power and significance level. The developed methodology for the Minimum Detectable Deformation is compelling for detecting anomalous displacements since the measured displacement can be substantiated with a probability.

We conclude that both detection strategies improve the early detection of impending sinkholes. We concluded that the spatio-temporal strategy is optimal for detecting a sinkhole size range and provides a general location of the possible impending sinkhole. The arc-based temporal strategy is very promising in early detecting of impending sinkholes. In this case, we found that the strategy detects an impending sinkhole after 3 months by setting a threshold. High values for the Detectability Power substantiate the results.

6

Recommendations

We took a careful look at the surface expression, datasets and two strategies. We showed that the surface expression is a viable approximation of the surface expression of an actual impending sinkhole and that the two strategies can detect these impending sinkholes. In this section we will talk about three recommendations. The first recommendation is the implementation of the strategies on a large subset. We tested the strategies on a small subset but now we want to monitor a larger area. Each strategy has its own limitations and this our thoughts are discussed in Section 6.1. The next recommendation will discuss the next step to test the strategies in Section 6.2. The last recommendation will discuss how to implement the two discussed strategies in Section 6.3.

6.1. Large scale Implementation

We want an early detection system to monitor a larger region than our used subset. Here, we discuss some thoughts on how to best scale the strategies to a larger area.

The spatio-temporal strategy based on the kinematic model is the easiest of the two to scale to a large area. One simply has to enlarge the grid. Making sure the grid covers the extent of the monitored area is sufficient for this strategy. However, expanding the grid will cost computationally.

The arc-based strategy is a little more trickier to scale-up. The arc-based strategy works best with small datasets. Larger datasets will exponentially make more arcs and these arcs are more prone to random flags. The large amount of arcs can be very computationally expensive and might clutter the results.

We recommend to implement this strategy subset-by-subset. The arc-based strategy works best on datasets containing ~60 points. We suggest implementing a grid onto the dataset whereby each posting can create a subset of around 60 points. We can quickly create subsets using the K-D tree data structure (Section 3.2.2). Processing a subset requires the steps: (1) arc creation and processing, (2) Arc characterization, (3) arc flagging, and (4) total flagged value computation. The last step is least optimized and will take the longest.

The current subset is 250 meters in the East direction and 400 meters in the North direction. The full dataset for track 88 has the dimension of 180 kilometers East to 317 kilometers North. We can make 4848 subsets out of the complete datasets with dimensions 250 by 400 meters. If we need 60 seconds per subset (i.e., subset creation, flagging), we need 81 hours of processing time. Eighty-one hours is too long. If we could cut down the number of subsets by a quarter, we have 1212 subsets that will need 20 hours of processing, still on the long side. We preferably want the subsets to overlap so that the edge observations can also lie closer to the center, making them more reliable. Doing this will again increase the number of subsets.

Therefore, implementing the arc-based strategy on the whole area might pose some difficulties. This method might be more applicable for monitoring small areas such as bridges, buildings, dams and ancient abandoned shafts.

6.2. Further Research on Impending Sinkhole Detection

The strategies were applied on a subset containing an artificial sinkhole. We want to discover actual sinkholes and a simulation might not take certain natural aspects into account (e.g., local geology). Therefore, we recommend to apply these strategies to actual collapsed sinkholes. One such an example of an area littered with collapsed sinkholes is given in the paper of Malinowska et al. (2019). Here, another ex post facto study was performed suggesting that PSI estimates contained displacement data indicative of an impending sinkhole. It would be very interesting if the proposed strategies would detect the impending sinkholes in time.

6.3. Implementation of the strategy

We investigated two strategies and how they improve early detection. Here, we discuss some recommendations on how to apply these strategies.

The spatio-temporal strategy provide a lot of information. It will not only give a rough estimate of the location but also provides an indication of its size. The provided information might be interesting for damage to the built environment. However, we found out that it was not the best strategy for early detection. The arc-based strategy detected the anomaly earlier. Therefore, we recommend apply this strategy for detecting phenomena location and extent.

The second strategy is the arc-based strategy. This strategy is better at detecting anomalous behavior in point measurements. However, this strategy will not provide any additional information. It simply marks a measurement point when it detects anomalous behavior. Thus, after the detection we still have to look for the location and size of the impending sinkhole.

We recommend to use the arc-based strategy for early detection and the spatio-temporal method for retrieve additional information. The optimal way would be to combine both strategies. For example, the arc-based strategy detects anomalous behavior where than the spatio-temporal strategy can be deployed for a more precise estimation of the location and the size.

Bibliography

Al-Halbouni, D., Holohan, E. P., Taheri, A., Watson, R. A., Polom, U., Schöpfer, M. P., Emam, S., & Dahm, T. (2019). Distinct element geomechanical modelling of the formation of sinkhole clusters within large-scale karstic depressions. *Solid Earth*, 10(4), 1219–1241.

Arkin, Y. & Gilat, A. (2000). Dead Sea sinkholes - an ever-developing hazard. *Environmental Geology*, 39(7), 711–722.

Atzori, S., Baer, G., Antonioli, A., & Salvi, S. (2015). InSAR-based modeling and analysis of sinkholes along the Dead Sea coastline. *Geophysical Research Letters*, 42(20), 8383–8390.

Augarde, C. E., Lyamin, A. V., & Sloan, S. W. (2003). Prediction of undrained sinkhole collapse. *Journal of Geotechnical and Geoenvironmental Engineering*, 129(3), 197–205.

Baer, G., Magen, Y., Nof, R. N., Raz, E., Lyakhovsky, V., & Shalev, E. (2018). InSAR measurements and viscoelastic modeling of sinkhole precursory subsidence: Implications for sinkhole formation, early warning, and sediment properties. *Journal of Geophysical Research: Earth Surface*, 123(4), 678–693.

Bamler, R. & Hartl, P. (1998). Synthetic aperture radar interferometry. *Inverse problems*, 14(4), R1.

Bekendam, R. & Pottgens, J. (1995). Ground movements over the coal mines of southern Limburg, The Netherlands, and their relation to rising mine waters. *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, 234, 3–12.

Bezuidenhout, C. & Enslin, J. (1970). Surface subsidence and sinkholes in the dolomitic areas of the Far West Rand, Transvaal, Republic of South Africa.

British Geological Survey (2020). Sinkholes (or dolines). https://www.bgs.ac.uk/caves/sinkholes/home.html.

Bruna, M. F. (2020). Analyzing subsidence in the Netherlands with attribute-enriched InSAR data. Master's thesis, TU Delft.

Buchignani, V., Avanzi, G. D., Giannecchini, R., & Puccinelli, A. (2007). Evaporite karst and sinkholes: a synthesis on the case of Camaiore (Italy). *Environmental Geology*, 53(5), 1037–1044.

Chang, L., Dollevoet, R., & Hanssen, R. (2014). Railway infrastructure monitoring using satellite radar data. *Int. J. Railw. Technol*, 3, 79–91.

Chang, L. & Hanssen, R. F. (2014). Detection of cavity migration and sinkhole risk using radar interferometric time series. *Remote Sensing of Environment*, 147, 56–64.

Cuenca, M. C., Hooper, A. J., & Hanssen, R. F. (2013). Surface deformation induced by water influx in the abandoned coal mines in Limburg, The Netherlands observed by satellite radar interferometry. *Journal of Applied Geophysics*, 88, 1–11.

Delaney, P. & McTigue, D. (1994). Volume of magma accumulation or withdrawal estimated from surface uplift or subsidence, with application to the 1960 collapse of Kilauea Volcano. *Bulletin of Volcanology*, 56(6), 417–424.

Denys, F. (2019). Former coal mining in the south of Limburg. Annual symposium on the theme "Post-mining risk management in the Netherlands".

Dunford, R., Su, Q., & Tamang, E. (2014). The pareto principle.

Dzurisin, D. (2006). *Volcano deformation: new geodetic monitoring techniques*. Springer Science & Business Media.

Eckardt, A. (1913). Die mechanischen einwirkungen des abbaus auf das verhalten des gebirges,. *Gluckauf 49*, 353/61, 397/403.

Economon, T. D., Palacios, F., Copeland, S. R., Lukaczyk, T. W., & Alonso, J. J. (2015). Su2: An open-source suite for multiphysics simulation and design. *Aiaa Journal*, 54(3), 828–846.

Ege, J. R. (1984). Mechanisms of surface subsidence resulting from solution extraction of salt. *Man-induced land subsidence: Geological Society of America Reviews in Engineering Geology*, 6, 203–221.

Elnabwy, M. T., Kaloop, M. R., & Elbeltagi, E. (2013). Talkha steel highway bridge monitoring and movement identification using RTK-GPS technique. *Measurement*, 46(10), 4282–4292.

Fell, R. & Hartford, D. (2018). Landslide risk management. In *Landslide Risk Assessment* (pp. 51–109). Routledge.

Ferretti, A., Prati, C., & Rocca, F. (2001). Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 39(1), 8–20.

Friedman, J. H., Bentley, J. L., & Finkel, R. A. (1977). An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software (TOMS)*, 3(3), 209–226.

Frumkin, A. & Raz, E. (2001). Collapse and subsidence associated with salt karstification along the Dead Sea. *Carbonates and Evaporites*, 16(2), 117–130.

Goel, K. & Adam, N. (2013). A distributed scatterer interferometry approach for precision monitoring of known surface deformation phenomena. *IEEE Transactions on Geoscience and Remote Sensing*, 52(9), 5454–5468.

Goudie, A. S. (2018). *Human impact on the natural environment*. John Wiley & Sons.

Gutiérrez, F., Benito-Calvo, A., Carbonel, D., Desir, G., Sevil, J., Guerrero, J., Martínez-Fernández, A., Karamplaglidis, T., García-Arnay, Á., & Fabregat, I. (2019). Review on sinkhole monitoring and performance of remediation measures by high-precision leveling and terrestrial laser scanner in the salt karst of the Ebro Valley, Spain. *Engineering Geology*, 248, 283–308.

Gutiérrez, F., Cooper, A. H., & Johnson, K. S. (2007). Identification, prediction, and mitigation of sinkhole hazards in evaporite karst areas. *Environmental Geology*, 53(5), 1007–1022.

Gutiérrez, F., Guerrero, J., & Lucha, P. (2008). A genetic classification of sinkholes illustrated from evaporite paleokarst exposures in Spain. *Environmental Geology*, 53(5), 993–1006.

Gutiérrez, F., Parise, M., Waele, J. D., & Jourde, H. (2014). A review on natural and human-induced geohazards and impacts in karst. *Earth-Science Reviews*, 138, 61–88.

Hanssen, R. F. (2001). *Radar interferometry: data interpretation and error analysis*, volume 2. Springer Science & Business Media.

Hanssen, R. F., van Leijen, F. J., Bruna, M. F. D., Li, Z., Spaans, J., Pepels, T., Houtepen, M., van der Sleen, V., & Bekendam, R. F. (2020). *Een monitorings- en risicosignaleringssysteem voorbodembeweging in Limburg*. Technical report. MRSS.

Heath, M. T. (2018). *Scientific Computing: An Introductory Survey, Revised Second Edition*. SIAM.

Heitfeld, M., Klunker, J., Rosner, P., Rosin, D., Spaans, J., & Rauwers, M. (2016). *Na-ijlende gevolgen steenkolenwinning Zuid-Limburg; Summary report with integrated Bow-Tie-Analysis*. Technical report.

Hermosilla, R. G. (2012). The Guatemala city sinkhole collapses. *Carbonates and evaporites*, 27(2), 103–107.

Intrieri, E., Gigli, G., Nocentini, M., Lombardi, L., Mugnai, F., Fidolini, F., & Casagli, N. (2015). Sinkhole monitoring and early warning: An experimental and successful GB-InSAR application. *Geomorphology*, 241, 304–314.

Johnson, K. S. (1989). Development of the Wink Sink in west Texas, USA, due to salt dissolution and collapse. *Environmental Geology and Water Sciences*, 14(2), 81–92.

Jones, C. E. & Blom, R. G. (2014). Bayou corne, louisiana, sinkhole: Precursory deformation measured by radar interferometry. *Geology*, 42(2), 111–114.

Kambezidis, H. (2012). 3.02 - the solar resource. In A. Sayigh (Ed.), *Comprehensive Renewable Energy* (pp. 27–84). Oxford: Elsevier.

Kampes, B. M. (2006). *Radar interferometry*. Springer.

Kim, J.-W., Lu, Z., & Degrandpre, K. (2016). Ongoing deformation of sinkholes in Wink, Texas, observed by time-series Sentinel-1A SAR Interferometry (Preliminary Results). *Remote Sensing*, 8(4), 313.

Kim, J.-W., Lu, Z., & Kaufmann, J. (2019). Evolution of sinkholes over Wink, Texas, observed by high-resolution optical and SAR imagery. *Remote Sensing of Environment*, 222, 119–132.

Kiyoo, M. (1958). Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Earthq Res Inst*, 36, 99–134.

Klünker, J. (2019). Dealing with post-mining-hazards from theoretical approach to practical remediation. Annual symposium on the theme "Post-mining risk management in the Netherlands".

Kotyrba, A. & Kortas, Ł. (2016). Sinkhole hazard assessment in the area of abandoned mining shaft basing on microgravity survey and modelling — case study from the upper silesia coal basin in poland. *Journal of Applied Geophysics*, 130, 62–70.

Kratzsch, H. (1983). *Mining subsidence engineering*. Springer Science & Business Media.

Luu, L.-H., Noury, G., Benseghier, Z., & Philippe, P. (2019). Hydro-mechanical modeling of sinkhole occurrence processes in covered karst terrains during a flood. *Engineering Geology*, 260, 105249.

López Dekker, P., de Zeeuw-van Dalfsen, E., R.F., H., & Heuff, F. (2018). Lecture notes in Geodesy and Natural Hazards, Module 1: SAR data basics and InSAR processing steps.

Malinowska, A. A., Witkowski, W. T., Hejmanowski, R., Chang, L., van Leijen, F. J., & Hanssen, R. F. (2019). Sinkhole occurrence monitoring over shallow abandoned coal mines with satellite-based persistent scatterer interferometry. *Engineering Geology*, 262, 105336.

Maneewongvatana, S. & Mount, D. M. (1999). It' s okay to be skinny, if your friends are fat. In *Center for geometric computing 4th annual workshop on computational geometry*, volume 2 (pp. 1–8).

Mogi, K. (1962). Study of elastic shocks caused by the fracture of heterogeneous materials and its relations to earthquake phenomena. *Bulletin of the Earthquake Research Institute, University of Tokyo (東京大學地震研究所)*, 40(1), 125–173.

Nof, R. N., Abelson, M., Raz, E., Magen, Y., Atzori, S., Salvi, S., & Baer, G. (2019). SAR interferometry for sinkhole early warning and susceptibility assessment along the dead sea, israel. *Remote Sensing*, 11(1), 89.

Nof, R. N., Baer, G., Ziv, A., Raz, E., Atzori, S., & Salvi, S. (2013). Sinkhole precursors along the dead sea, israel, revealed by sar interferometry. *Geology*, 41(9), 1019–1022.

Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the seismological society of America*, 75(4), 1135–1154.

Paine, J. G., Buckley, S. M., Collins, E. W., & Wilson, C. R. (2012). Assessing collapse risk in evaporite sinkhole-prone areas using microgravimetry and radar interferometry. *Journal of Environmental and Engineering Geophysics*, 17(2), 75–87.

Panton, R. L. (2006). *Incompressible flow*. John Wiley & Sons.

Pöttgens, J. (1985). Bodemhebung durch ansteigen- des grubenwasser [uplift as a result of rising mine wa- ters]. *The Development Science and Art of Minerals Surveying, Proceedings VIth International Congress, International Society for Mine Surveying, Harrogate, United Kingdom*, 2, 928–938.

Redactie Algemeen Nederlands Persbureau (2020). Gat kerkrade vermoedelijk veroorzaakt door verzakte eeuwenoude mijnschacht. https://www.hartvannederland.nl/nieuws/2020/groot-sinkhole-kerkrade-omwonenden-geevacueerd/.

Redactie RTL Nieuws (2020). Kademuur ingestort in centrum amsterdam, vermoedelijk door sinkhole. https://www.rtlnieuws.nl/nieuws/nederland/artikel/5180850/kade-ingestort-centrum-amsterdam.

Reddish, D. & Whittaker, B. (1989). *Subsidence: occurrence, prediction and control*. Elsevier.

Ren, G. & Li, J. (2008). A study of angle of draw in mining subsidence using numerical modeling techniques. *Electron. J. Geotech. Eng*, 13, 1–14.

Ren, G., Reddish, D., & Whittaker, B. (1987). Mining subsidence and displacement prediction using influence function methods. *Mining Science and Technology*, 5(1), 89–104.

Roest, J. (2019). The legacy of the coal mining in limburg. Annual symposium on the theme "Post-mining risk management in the Netherlands".

Shalev, E. & Lyakhovsky, V. (2012). Viscoelastic damage modeling of sinkhole formation. *Journal of Structural Geology*, 42, 163–170.

Skolnik, M. I. et al. (1962). *Introduction to radar systems*, volume 3. McGraw-hill New York.

Sparling, D. W. (2016). Chapter 12 - modeling in ecotoxicology. In D. W. Sparling (Ed.), *Ecotoxicology Essentials* (pp. 361–390). San Diego: Academic Press.

Teunissen, P. J. G. (2000). *Testing theory; an introduction*. Delft: Delft University Press, 1 edition.

Theron, A. & Engelbrecht, J. (2018). The role of earth observation, with a focus on SAR interferometry, for sinkhole hazard assessment. *Remote Sensing*, 10(10), 1506.

van Leijen, F. J. (2014). *Persistant Scatterer Interferometry based on geodetic estimation theory*. PhD thesis, TU Delft.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., & SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17, 261–272.

Waltham, T., Waltham, A. C., Bell, F. G., & Culshaw, M. G. (2005). *Sinkholes and subsidence: karst and cavernous rocks in engineering and construction*. Springer Science & Business Media.

Wang, J., Peng, X., & hui Xu, C. (2011). Coal mining GPS subsidence monitoring technology and its application. *Mining Science and Technology (China)*, 21(4), 463–467.

Whyatt, J. & Varley, F. (2008). Catastrophic failures of underground evaporite mines. In *Proceedings: 27th International Conference on Ground Control in Mining* (pp. 29–31).: Morgantown.

Williams, P. (2003). Encyclopedia of caves and karst science.

Zhu, W., Zhang, Q., Ding, X., Zhao, C., Yang, C., Qu, F., & Qu, W. (2014). Landslide monitoring by combining of cr-insar and gps techniques. *Advances in space research*, 53(3), 430–439.

A

Links to external content

This appendix acts as an index for the different notebooks, data, and scripts used to obtain the results for this thesis. Most items will be referring to either GitHub (https://github.com/maxfelius) or my personal website (max.felius.nl).

- Notebook, 2.2.2.Figure_Angle_Definitions, https://github.com/maxfelius/Notebooks_List/blob/main/2.2.2.Figure_Angle_Definitions.ipynb
- Notebook, 2.2.2.Selected geometric models (influence functions) https://github.com/maxfelius/Notebooks_List/blob/main/2.2.2.Selected%20geometric%20models%20(influence%20functions).ipynb
- Notebook, 3.1.1.Comparison Gaussian Influence Function with Normal Distribution https://github.com/maxfelius/Notebooks_List/blob/main/3.1.1.Comparison%20Gaussian%20Influence%20Function%20with%20Normal%20Distribution.ipynb
- Notebook, 3.2.3.Precision_Deformation_Time_Series https://github.com/maxfelius/Notebooks_List/blob/main/3.2.3.Precision_Deformation_Time_Series.ipynb
- Notebook, 3.2.4.MDD_DP_Figure https://github.com/maxfelius/Notebooks_List/blob/main/3.2.4.MDD_DP_Figure.ipynb
- Notebook, 4.1.1.Ground_Movement https://github.com/maxfelius/Notebooks_List/blob/main/4.1.1.Ground_Movement.ipynb
- Notebook, 4.1.1.Ground_Movement_Over_Time https://github.com/maxfelius/Notebooks_List/blob/main/4.1.1.Ground_Movement_Over_Time.ipynb
- Notebook, 4.1.2.Comparison_Kinematic_Physical_Model https://github.com/maxfelius/Notebooks_List/blob/main/4.1.2.Comparison_Kinematic_Physical_Model.ipynb
- Folder, 4.1.3.Real-world cases https://github.com/maxfelius/Notebooks_List/tree/main/4.1.3.Real-world%20cases
- Notebook, 4.2.1.Implement_Accelerating_Sinkhole_Into_Dataset https://github.com/maxfelius/Notebooks_List/blob/main/4.2.1.Implement_Accelerating_Sinkhole_Into_Dataset.ipynb

- Notebook, 4.2.1.Implement_Linear_Sinkhole_Into_Dataset
  https://github.com/maxfelius/Notebooks_List/blob/main/4.2.1.Implement_Linear_Sinkhole_Into_Dataset.ipynb
- Notebook, 4.2.2.Histogram_MDS
  https://github.com/maxfelius/Notebooks_List/blob/main/4.2.2.Histogram_MDS.ipynb
- Notebook, 4.2.2.Methodology_Detectability_Map_v2
  https://github.com/maxfelius/Notebooks_List/blob/main/4.2.2.Methodology_Detectability_Map_v2.ipynb
- Notebook, 4.2.3.Maps_MDD_and_DP
  https://github.com/maxfelius/Notebooks_List/blob/main/4.2.3.Maps_MDD_and_DP.ipynb
- Folder, 4.3.1.Spatio-temporal Strategy based on the kinematic model
  https://github.com/maxfelius/Notebooks_List/tree/main/4.3.1.Spatio-temporal%20Strategy%20based%20on%20the%20kinematic%20model
- Folder, 4.4.1.Arc Characterization
  https://github.com/maxfelius/Notebooks_List/tree/main/4.4.1.Arc%20Characterization
- Folder, 4.4.2.Temporal Strategy based on arcs
  https://github.com/maxfelius/Notebooks_List/tree/main/4.4.2.Temporal%20Strategy%20based%20on%20arcs

A.1. Sinkhole Package Scripts

In this section the scripts are listed which are part of the sinkhole package used in multiple notebooks. The notes in the scripts will explain the workings.

- Python Script, arc_flagging.py,
  https://github.com/maxfelius/sinkhole_scripts/blob/main/arc_flagging.py
- Python Script, create_tuple_list.py,
  https://github.com/maxfelius/sinkhole_scripts/blob/main/create_tuple_list.py
- Python Script, deformation_analysis.py,
  https://github.com/maxfelius/sinkhole_scripts/blob/main/deformation_analysis.py
- Python Script, detectability_map.py,
  https://github.com/maxfelius/sinkhole_scripts/blob/main/detectability_map.py
- Python Script, functions.py,
  https://github.com/maxfelius/sinkhole_scripts/blob/main/functions.py
- Python Script, geometric_models.py,
  https://github.com/maxfelius/sinkhole_scripts/blob/main/geometric_models.py
- Python Script, get_delta_days.py,
  https://github.com/maxfelius/sinkhole_scripts/blob/main/get_delta_days.py

- Python Script, influence_function.py, https://github.com/maxfelius/sinkhole_scripts/blob/main/influence_function.py
- Python Script, linear_model.py, https://github.com/maxfelius/sinkhole_scripts/blob/main/linear_model.py
- Python Script, mogi_model.py, https://github.com/maxfelius/sinkhole_scripts/blob/main/mogi_model.py
- Python Script, process_arc_connections.py, https://github.com/maxfelius/sinkhole_scripts/blob/main/process_arc_connections.py
- Python Script, process_arc_flagging.py, https://github.com/maxfelius/sinkhole_scripts/blob/main/process_arc_flagging.py
- Python Script, spatio_temporal_functions.py, https://github.com/maxfelius/sinkhole_scripts/blob/main/spatio_temporal_functions.py

B

Arc-based Strategy Output Table

Table B.1: Table of the results from the arc-based strategy with a threshold set to 150. All the marked points with a Total Flagged Value above the threshold (150) are shown in the table. It clearly shows that the point with ID 82 is marked from 14 September till the end of the time series.

Epoch Date	ID	pnt_rdx	pnt_rdy	Total Flagged Value
d_20180318	17	202994	318388	174.0
d_20180803	31	202875	318394	163.0
d_20180914	82	202944	318351	179.0
d_20180926	82	202944	318351	202.0
d_20181002	74	202940	318336	163.0
d_20181002	82	202944	318351	225.0
d_20181002	75	202942	318336	171.0
d_20181008	74	202940	318336	155.0
d_20181008	82	202944	318351	228.0
d_20181008	75	202942	318336	183.0
d_20181026	82	202944	318351	233.0
d_20181026	75	202942	318336	188.0
d_20181101	82	202944	318351	232.0
d_20181101	75	202942	318336	161.0
d_20181119	82	202944	318351	235.0
d_20181119	75	202942	318336	164.0
d_20181201	74	202940	318336	151.0
d_20181201	82	202944	318351	238.0
d_20181201	75	202942	318336	178.0
d_20181207	10	202874	318366	157.0
d_20181207	82	202944	318351	228.0
d_20181213	10	202874	318366	165.0
d_20181213	82	202944	318351	231.0
d_20181213	75	202942	318336	164.0
d_20181213	36	202880	318410	151.0
d_20181219	74	202940	318336	152.0
d_20181219	82	202944	318351	234.0
d_20181219	75	202942	318336	181.0
d_20181219	36	202880	318410	151.0
d_20181225	10	202874	318366	158.0
d_20181225	74	202940	318336	189.0
d_20181225	82	202944	318351	237.0
d_20181225	75	202942	318336	202.0
d_20181225	36	202880	318410	155.0
d_20181231	10	202874	318366	173.0
d_20181231	74	202940	318336	227.0
d_20181231	82	202944	318351	240.0
d_20181231	75	202942	318336	219.0
d_20181231	36	202880	318410	169.0