



## Huidige en toekomstige kosten MT/LT warmtenetten

### Thema 2 – Project C

#### Hoofdauteur

Lieke Hüsken *Deltares*

17 november 2022

## Huidige en toekomstige kosten MT/LT warmtenetten



Deltares  
17 november 2022

Auteur: Lieke Hüsken *Deltares*  
i.s.m. projectteam van Eneco, Vattenfall, HVC, ENnatuurlijk, SV Purmerend, Firan, Enpuls, Heijmans

Kwaliteitsborging: Ivo Pothof *Deltares*

Dit project is uitgevoerd als onderdeel van het Innovatieplan WarmingUP. Dit is mede mogelijk gemaakt door subsidie van de Rijksdienst voor Ondernemend Nederland (RVO) in het kader van de subsidieregeling Meerjarige Missiegedreven Innovatie Programma's (MMIP), bij RVO bekend onder projectnummer TEUE819001. WarmingUP geeft invulling aan MMIP-4 – Duurzame warmte en koude in gebouwde omgeving en levert daarmee een bijdrage aan Missie B – Een CO<sub>2</sub>-vrije gebouwde omgeving in 2050.

[Projectnummer](#)  
11205151

[Keywords](#)  
Aanlegkosten, uitvoeringskosten, kostendrijvers warmtenetten, gebouwde omgeving, kentallen, multivariate regressie

[Jaar van publicatie](#)  
2022

[Meer informatie](#)  
Ivo Pothof  
T 06 2288 3518  
E Ivo.Pothof@Deltares.nl

Maart 2022 ©

Alle rechten voorbehouden. Niets uit deze uitgave mag worden veeveelvoudigd, opgeslagen in een geautomatiseerd gegevens bestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

# Voorwoord, disclaimer

Dit onderzoek is uitgevoerd door Deltares in nauwe samenwerking met een projectteam van warmtebedrijven, netbeheerders en aannemers. Ondanks het feit dat praktisch het hele project is uitgevoerd met Corona-bependingen (alleen online meetings), is toch een goede samenwerking tot stand gekomen. Het projectteam heeft een bruikbaar dataformaat opgesteld om data over de aanleg en aanlegkosten van warmteleidingen eenduidig te verzamelen. De tussenresultaten zijn verschillende keren besproken met het projectteam om de relevante inzichten en resultaten zo goed mogelijk uit de verzamelde dataset te halen. We bedanken het projectteam voor de constructieve samenwerking in dit project.

De dataset bestaat uit slechts 63 projecten. Bovendien hebben de energietransitie en de geopolitieke ontwikkelingen grote impact op de grondstofprijzen en beschikbare operationele capaciteit voor de aanleg van warmtenetten. Deze factoren zijn daarmee onvoldoende om alle verbanden betrouwbaar te kwantificeren, maar voldoende om inzicht te krijgen in de belangrijkste (omgevings)factoren die bepalend zijn voor de aanlegkosten van terreinleidingen van warmtenetten. Hierdoor is het opgestelde model niet geschikt voor het doorrekenen van aanlegkosten in willekeurige woonwijken tbv transitieplannen of business cases. De ontwikkelde methodiek is wél bruikbaar gebleken, maar met het grote aantal relevante omgevingsfactoren is een veel grotere dataset nodig en minder volatiliteit in de toekomstige grondstofprijzen en personele capaciteit.

Het opgestelde model kan verder gebruikt worden als een communicatiemiddel met gemeenten, provincies, woningcorporaties of andere stakeholders en laat zien welke omgevingsfactoren van invloed zijn op de aanlegkosten in een specifiek woonwijk. Hiermee kan de communicatie, het draagvlak en begrip vergroot worden voor de realisatie van een warmtenet.

Dank aan alle individuen en organisaties die hieraan hebben bijgedragen: Martijn Matijssen (Vattenfall), Niels van Schie (Eneco), Richard van Ballegooijen (Ennatuurlijk), Patrick van Duren (Firan), Berry de Jong (SVP), René Frinks (Heijmans), Dennis van Erp (HVC groep), Jonathan Nuttall (Deltares), Lieke van der Most (Deltares), Amine Aboufirass (Deltares), Sonja Wanke (Deltares), Lőrinc Mészáros (Deltares), Jorn Mieras (stagiaire Deltares).

Lieke Hüsken en Ivo Pothof

# Inhoudsopgave

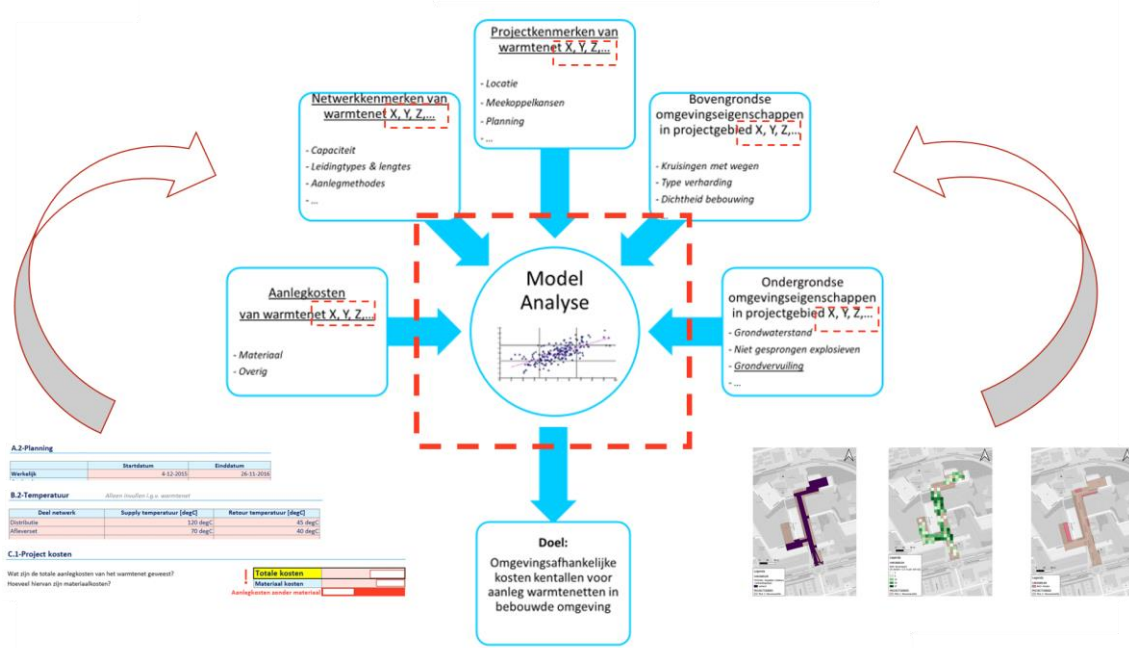
<b>Voorwoord, disclaimer</b>	<b>3</b>
<b>Samenvatting</b>	<b>6</b>
<b>1 Hoofdstuk</b>	<b>8</b>
1.1 Doel en urgentie	9
1.2 Projectaanpak	9
1.3 Onderzoeksopbouw	10
1.4 Deliverables	12
<b>2 Samenvatting literatuurstudie</b>	<b>13</b>
2.1 Technieken voor kostenramingen	13
2.2 Statistische modellen voor kostenramingen van warmtenetten	14
2.3 Omgevingsfactoren in energie transitie modellen	14
2.4 Kostendrijvers	17
<b>3 Opstellen database</b>	<b>18</b>
3.1 Definiëren databehoeftes en databeschikbaarheid	18
3.2 Ontwikkelen template	23
3.3 GIS data	24
3.4 Opschonen Database	25
3.5 Omschrijving Database	26
<b>4 Data Analyse</b>	<b>27</b>
4.1 Data beschouwing	27
4.1.1 Aanlegkosten (totaal en per meter)	27
4.1.2 Projecteigenschappen	31
4.1.2.1 Projecteigenschap “Bedrijf”	32
4.1.3 Technische eigenschappen	35
4.1.4 Omgevingseigenschappen	38
4.2 Correlatie analyse	44
4.3 Modelanalyse – Multivariate regressie	45
4.3.1 Model: Uitvoeringskosten	46
4.3.2 Model: M_Uitvoeringskosten	50
4.3.3 Model: Baseline	53
<b>5 Conclusies en aanbevelingen</b>	<b>57</b>
5.1 Conclusies	57
5.1.1 Relevante omgevingseigenschappen	57
5.1.2 Relevante project- en technische eigenschappen	58

5.1.3	Kostenmodel	58
5.1.4	Baseline modellen	59
5.2	Aanbevelingen	59
5.2.1	Postcode en vergunningsvoorwaarden	59
5.2.2	Valideren resultaten data analyse in de praktijk	59
5.2.3	Valideren resultaten data analyse met grotere database	59
5.2.4	Synergie met energie transitie rekenmodellen (vervolg)	60
5.2.5	Aandachtpunten voor data inventarisatie	60
5.2.6	Ontbrekende GIS data	61
<b>Bibliography</b>		<b>62</b>
<b>Bijlagen</b>		<b>63</b>
A.	Longlist Variabelen - inhoud en bronnen	63
	Projecteigenschappen	63
	Technische eigenschappen	64
	Omgevingseigenschappen	66
	Aanlegkosten	68
B.	Template data inventarisatie warmtebedrijven	71
C.	Gevoeligheidsanalyse grootte projectzone	75
D.	Model resultaten - Uitvoeringskosten	78
E.	Model resultaten - M_Uitvoeringskosten	81

# Samenvatting

Het project *Huidige en toekomstige kosten MT/LT warmtenetten* is onderdeel van thema 2 van het innovatieprogramma WarmingUp. WarmingUp richt zich op het ontwikkelen van technische en sociaal-maatschappelijke kennis met als doel daar praktisch bruikbare instrumenten van te maken voor het ontwerpen en realiseren van kosteneffectieve, maatschappelijk aanvaardbare, betrouwbare, collectieve warmtesystemen in de gebouwde omgeving die door duurzame warmtebronnen worden gevoed. In *Project 2C: Huidige en toekomstige kosten MT/LT warmtenetten* wordt er een database opgezet om kwantitatief te verkennen welke factoren, naast de omvang en technische eigenschappen van het netwerk, invloed hebben op de aanlegkosten van warmtenetten.

Aan de basis van dit project ligt een kwantitatieve aanpak, aangevuld en vormgegeven door gebruik te maken van praktijkkennis. In deze kwantitatieve aanpak worden de kosten van historische projecten (in het verleden uitgevoerde en afgeronde projecten) gecombineerd met publieke GIS-data om te onderzoeken welke (omgevings)factoren de aanlegkosten significant beïnvloeden (Figuur 1-1). De scope van de verzamelde projectkosten is uitsluitend gericht op de uitvoeringsfase met daadwerkelijke aanleg van terreinleidingen en direct omgevingsmanagement in de bestaande gebouwde omgeving.



Figuur 1-1: Schematische weergave van de aanpak

De database bevat 63 projecten, uitgevoerd tussen 2015 en 2020. Elk project heeft 13 technische eigenschappen en 20 omgevingseigenschappen. Hiermee zijn 2 modellen opgesteld. Een voor de uitvoeringskosten en een voor de uitvoeringskosten per meter warmteleiding. Deze gemodelleerde uitvoeringskosten zijn exclusief materiaalkosten, omdat de materiaalkosten voor een groot aantal projecten niet aangeleverd konden worden. De belangrijkste verklarende variabelen voor de uitvoeringskosten zijn het postcode-gebied, de lengte van het warmtenet, het aantal aansluitingen,

de aanwezigheid van waterlichamen in het projectgebied, de wegbreedte, het percentage wegen in het projectgebied en het gemiddelde bouwjaar van de woningen. De onderzochte database is te klein om voldoende statistische onderbouwing te leveren. Echter, bovenstaande bevindingen wijzen er wel op dat er een rol is voor omgevingsfactoren in hun bijdrage aan het voorspellen van de uitvoeringskosten van warmtenetten. De regressiemodellen presteren beter dan het baseline model (VESTA MAIS).

Door het projectteam is ervaren hoe waardevol het is om over kwantitatieve data te beschikken en als basis te gebruiken voor discussies. Er wordt dan ook door het projectteam aanbevolen om de huidige database te bewaren en gestructureerd uit te breiden. Een grotere omvang van de beschikbare database kan leiden tot waardevolle inzichten op gebied van kosteninschattingen maar kan daarmee ook richting geven aan (het monitoren van) kostenbesparende innovaties. Daarnaast is het wenselijk om een grotere diversiteit aan postcodes, bedrijven en groottes (in lengte) van projecten in de dataset op te nemen.

# 1 Hoofdstuk

Het project *Huidige en toekomstige kosten MT/LT warmtenetten* is onderdeel van thema 2 van het innovatieprogramma WarmingUp. WarmingUp richt zich op het ontwikkelen van technische en sociaal-maatschappelijke kennis met als doel daar praktisch bruikbare instrumenten van te maken voor het ontwerpen en realiseren van kosteneffectieve, maatschappelijk aanvaardbare, betrouwbare, collectieve warmtesystemen in de gebouwde omgeving die door duurzame warmtebronnen worden gevoed.

WarmingUp is onderverdeeld in 7 thema's. In vier samenhangende projecten worden er in Thema 2: Grootschalige en kosteneffectieve aanleg van warmtenetten kennis en technologieën ontwikkeld om het opschalen van leidingnetwerken op een kostenefficiënte manier mogelijk te maken. Door middel van verlaagde aanvoertemperaturen, slimme aanlegmethodes, inzicht in kostenkanten en faalkansmodellen worden er voor alle marktpelers laagdrempelige high-impact ontwikkelingen geboden.

In *Project 2C: Huidige en toekomstige kosten MT/LT warmtenetten* (zie Textbox 1.1 voor definities) wordt er een database opgezet om kwantitatief te verkennen welke factoren, naast de omvang en technische eigenschappen van het netwerk, invloed hebben op de aanlegkosten van warmtenetten. Denk daarbij bijvoorbeeld aan de beschikbare ruimte boven en onder de grond, de aanwezigheid van vervuilde grond, of bomen op of in de buurt van het tracé. De rapportage die voor u ligt betreft de eind rapportage van project 2C, waarin de afgesproken deliverables/resultaten worden gedeeld maar waarin ook het proces om hiertoe te komen is vastgelegd.

Textbox 1.1 Definities temperaturen warmtenetten (WarmingUp, 2020)

---

*In de hele keten van transport en distributie in het warmtenet tot aan de aflevering van warmte bij de klant komen verschillende temperatuurniveaus voor. Om het temperatuurniveau te duiden worden vaak benamingen gebruikt voor de aanvoertemperatuur in het net zoals hoge, midden en lage temperatuur. De interpretatie van deze benamingen is echter vaak verschillend en kan ook afhangen van de context waarin deze gebruikt wordt. In de huidige rapportage wordt de volgende definitie voor "LT / MT" gebruikt,*

- *Hoge temperatuur (HT)  $T_a > 70^\circ\text{C}$*
- ***Midden temperatuur (MT)  $55 \leq T_a \leq 70^\circ\text{C}$***
- ***Lage temperatuur (LT)  $25 \leq T_a < 55^\circ\text{C}$***
- *Zeer lage temperatuur (ZLT)  $T_a < 25^\circ\text{C}$*

*Deze definitie betreft de benadering vanuit het afleverpunt en de mogelijkheden die de geleverde aanvoertemperatuur biedt en aansluitend bij het perspectief van veel stakeholders.*

---



## 1.1 Doel en urgentie

Het doel van dit deelproject is ten eerste om **beter inzicht te krijgen in de factoren die in de bebouwde omgeving invloed hebben op de aanlegkosten van warmtenetten** en ten tweede om deze inzichten in te zetten om te komen tot **realistischere kosteninschattingen van warmtenet aanlegprojecten in de bebouwde omgeving**. Project 2C beoogd daarmee de *bestaande bandbreedte* van onzekerheid bij kosteninschattingen van warmtenet aanlegprojecten al *vroegtijdig* te verkleinen door rekening te houden met de invloed van omgevingsfactoren op de aanlegkosten. Onderliggend is daarmee de onderzoekshypothese dat omgevingsfactoren van significante invloed zijn op de aanlegkosten, en dat het niet meenemen van deze omgevingsfactoren bij kosteninschattingen een verklaring is voor van een deel van de bestaande bandbreedte van onzekerheid.

Met *bestaande bandbreedte* wordt bedoeld de onzekerheid die gepaard gaat met de momenteel gebruikte rekenmodellen (en bijbehorende kentallen), zoals het Vesta model, om kosten voor de aanleg van warmtenetten in te schatten. Momenteel ontbreekt het bouwers en gebruikers van energietransitie rekenmodellen, zoals bijvoorbeeld CEgoia en VESTA, aan een uniform overzicht van kentallen wat resulteert in uiteenlopende en afwijkende inschattingen van benodigde investeringen. In de praktijk blijkt soms de nodige investering 2 a 3 keer groter dan aan de voorkant verwacht. Om deze reden is het van belang om de bandbreedte van onzekerheid te verkleinen en te komen tot realistischer kosten-inschattingen

Met *vroegtijdig* wordt bedoeld al vroeg in het besluitvormingsproces, wanneer detailinformatie van een uitgewerkt technisch ontwerp (nog) niet beschikbaar is maar wanneer er al wel belangrijke besluiten genomen moeten worden, bijvoorbeeld in het kader van het opstellen van een lokale of regionale energie transitie strategie. De te verwachten kosten van de aanleg van warmtenetten is een belangrijk criterium tijdens het opstellen en afwegen van energie transitie strategieën.

Daarnaast kunnen de resultaten van dit onderzoek ook waardevolle input leveren voor andere strategische keuze momenten zoals bij begrotingsprocessen of ter overweging tijdens het opstellen van de aanbestedingsstrategie voor een warmtenet aanlegproject. Door de projectpartners is aangegeven dat de beoogde resultaten naar verwachting ook zullen leiden tot een betere projectstart doordat de verschillende verwachtingen van de betrokken partijen wat betreft aanlegkosten meer op een lijn zitten en al beter onderbouwd kunnen worden.

## 1.2 Projectaanpak

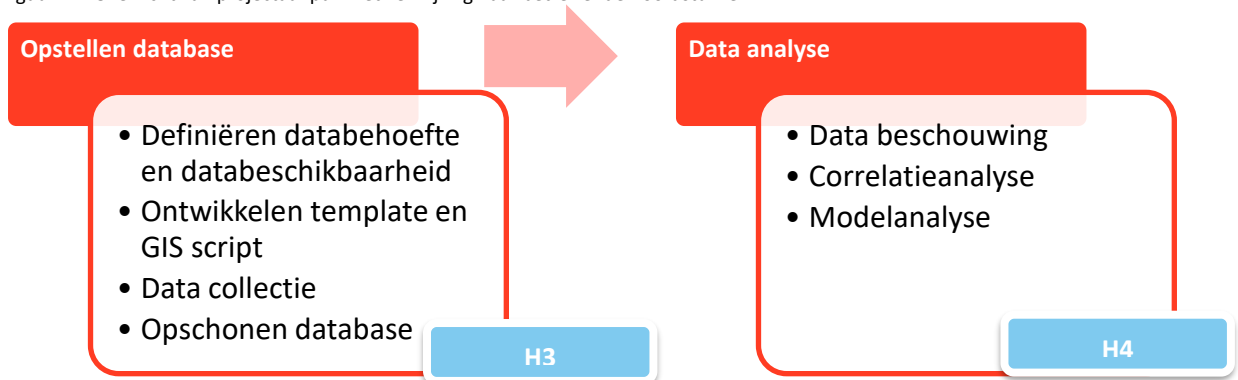
Aan de basis van dit project ligt een kwantitatieve aanpak, aangevuld en vormgegeven door gebruik te maken van praktijk kennis. In deze kwantitatieve aanpak worden de kosten van historische projecten (in het verleden uitgevoerde en afgeronde projecten) onderzocht. Specifiek wordt er onderzocht of er statistische verbanden gevonden kunnen worden tussen de projectkosten van de historische projecten en omgevingseigenschappen in de projectgebieden van de historische projecten.

Daarmee is het project opgebouwd uit 2 delen ( Figuur 1-1), hieronder kort omschreven. Verdere detaillering van de verschillende (deel)stappen zal in de betreffende hoofdstukken aan bod komen.

Het eerste deel betreft het vullen van een database met gegevens van historische warmtenet aanlegprojecten in de bebouwde omgeving. Om tot een database te komen zijn, in nauwe samenwerking met de projectpartners in het projectteam, verschillende stappen doorlopen. Allereerst het definiëren van de data behoefte (*welke data is er nodig om de onderzoeksvraag te beantwoorden?*) en de data beschikbaarheid (*is de benodigde data beschikbaar, en waar?*) Dit heeft geleid tot het identificeren van twee hoofdbronnen waar de benodigde data vandaan komt: data afkomstig van de projectpartners zelf en data afkomstig uit nationale GIS bestanden. De data afkomstig van de projectpartners wordt door middel van een gezamenlijk opgestelde template verzameld en de benodigde GIS data wordt geïmporteerd uit verschillende nationale data bestanden voor de betreffende projectgebieden. Wanneer de data eenmaal verzameld is worden de gegevens uit de twee hoofdbronnen samengevoegd en opgeschoond.

Wanneer de database gereed is kan deze geanalyseerd worden (deel 2). Dit gebeurt eerst door de data op verschillende manieren te beschouwen (*Denk bijvoorbeeld aan: hoe ziet de spreiding er uit, hoeveel data punten zijn er en zien we bepaalde patronen?*) Vervolgens wordt er een correlatie analyse uitgevoerd tussen de verschillende verklarende variabelen (explanatory variables) en de aanlegkosten als response variabele. Op basis van de resultaten uit de correlatie analyse, wederom aangevuld met belangrijk praktijk kennis voor de interpretatie van de resultaten, wordt vervolgens een kostenmodel opgezet.

Figuur 1-1 Overzicht van projectaanpak met verwijzing naar betreffende hoofdstukken



### 1.3 Onderzoeksop

Binnen project 2C is, in overleg met de projectpartners, een duidelijke afbakening van de onderzoeksop bepaald. Dit heeft met name betrekking op de te onderzoeken kosten componenten en delen van het warmtenetwerk.

**De uitvoeringsfase** – Er wordt alleen gekeken naar de kosten die gemaakt worden tijdens de uitvoeringsfase - de daadwerkelijke aanleg – van warmtenetten. Daar vallen materiaalkosten en civiele werkzaamheden onder, maar ook de bijkomende kosten zoals project en omgevingsmanagement, grondaankoop etc.

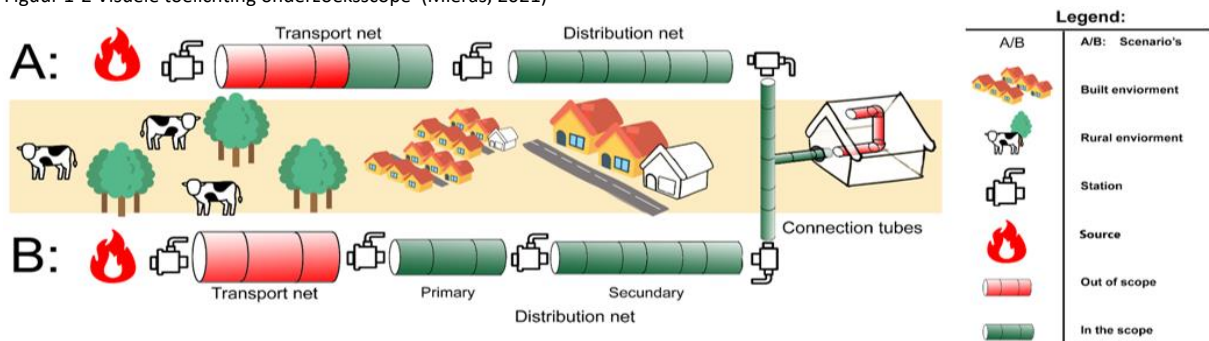
Kosten die gemaakt worden in de planning en ontwerpfasen van het project, voorafgaand aan de uitvoeringsfase, worden niet meegenomen. Deze kosten worden naar verwachting hoofdzakelijk beïnvloed door de warmtevraag en beschikbaarheid van bronnen en niet zo zeer de omgevingsfactoren.

- i. **Bestaande bouw** - Het project richt zich op de bestaande bebouwde omgeving. Er wordt gekeken naar het aanleggen van warmtenetten voor bestaande huizen die van het gas af

worden gehaald of inbreidingslocaties van nieuwbouw in bestaande bouw. We kijken niet naar nieuwbouw(wijken).

- ii. **“Terreinleidingen”** Het project richt zich op de aanlegkosten van leidinginfrastructuur (transport- en/of distributie- en/of, aansluitleidingen). Er wordt niet gekeken naar binnenpandige, woning afhankelijke aanlegkosten en er wordt uitgegaan van een aansluiting achter de voordeur met een meterkast op de begane grond (~3m naar binnen). We erkennen het belang en de grote variatie in deze woning afhankelijke installatiekosten maar dit valt buiten de scope van dit onderzoek. Ook de kosten van de aanleg van en aansluiting met de bronnen of (tussen) stations valt buiten de onderzoeksscope. Figuur 1-2 geeft deze scope visueel weer.

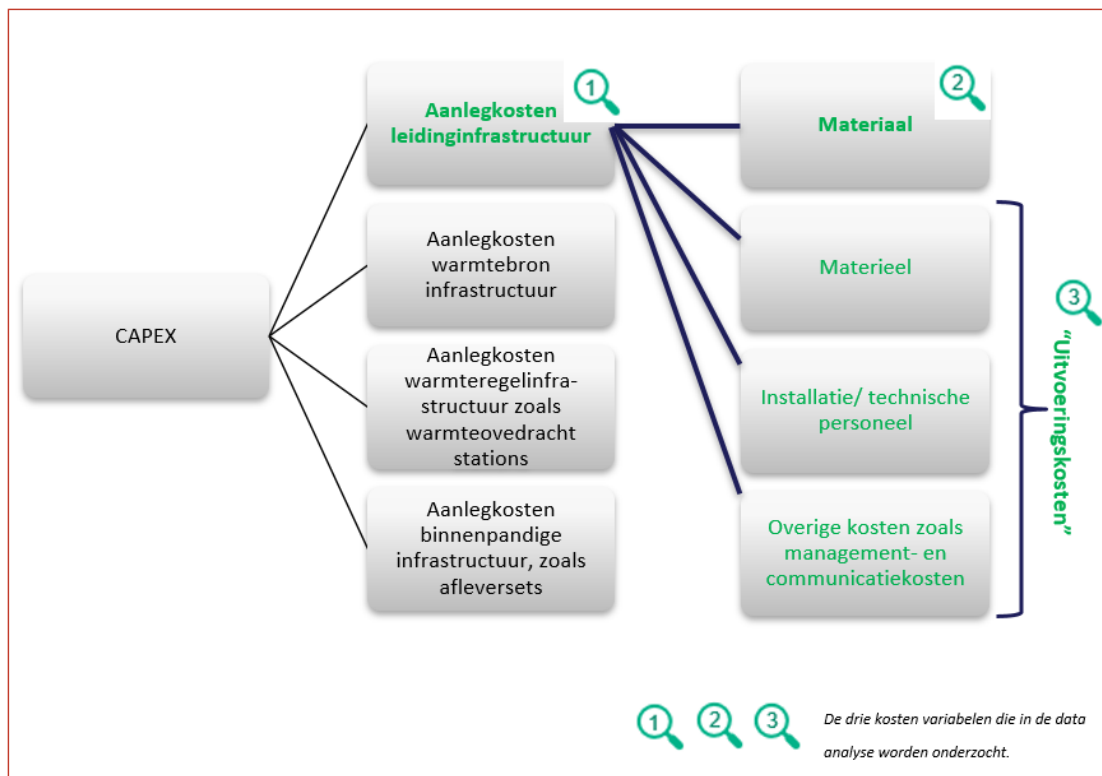
Figuur 1-2 Visuele toelichting onderzoeksscope (Mieras, 2021)



- iii. **Aanlegkosten leidinginfrastructuur** - Voor dit project zijn we vooral op zoek naar de omgevingsfactoren die de aanlegkosten van leidinginfrastructuur (significant) beïnvloeden. Figuur 1-3 geeft de opbouw weer van verschillende kostencomponenten die verzameld zijn in het project en hoe deze bijdragen aan de totale aanlegkosten die passen binnen de scope. De groen gedrukte onderdelen, gemarkeerd met een vergrootglas, zijn de kosten variabelen (response variabelen) die verzameld zijn. Te zien is dat de CAPEX (Capital Expenses) voor het aanleggen van een warmtenet opgedeeld is in verschillende componenten die betrekking hebben op verschillende onderdelen van het warmtenet. In dit project wordt enkel ingezoomd op het onderdeel “Aanlegkosten leidinginfrastructuur”(1) met de bijbehorende sub-posten die daar onder vallen.

Er is veel variatie tussen de bedrijven in welke uitgaven onder welke post geregistreerd worden. Zo beslaat bij bedrijf X de post materieel enkel de apparatuur die nodig is voor de aanleg terwijl bij bedrijf Y hier ook de benodigde specifieke personele kosten voor bedienen van de apparatuur en machines bij inbegrepen zit. “Materiaal” (2) is de enige kostenpost die op uniforme wijze aangeleverd kon worden door de partners. De overige niet, waardoor deze gebundeld zijn onder “uitvoeringskosten” (3). Het is niet uit te sluiten dat deze kostenpost ook kosten bevat die bijvoorbeeld voortkomen uit “aanlegkosten warmtebron infrastructuur” of een van de andere CAPEX onderdelen.

Figuur 1-3 Duiding beschouwde kosten



## 1.4 Deliverables

Binnen project 2C wordt toegewerkt naar verschillende (deel)resultaten en levert de volgende deliverables:

- ✓ **Een database** met gegevens, waaronder kosten, van historisch uitgevoerde warmtenet aanlegprojecten. Deze database is niet openbaar en valt onder de Non-Disclosure Agreement die is afgesproken met de warmtebedrijven.
  - *Voor de totstandkoming van de database is een template (Excel) en een script (python) ontwikkeld, deze zijn wel openbaar en kunnen opgevraagd worden bij Deltares.*
- ✓ **Een overzicht van omgevingseigenschappen** die de aanlegkosten van warmtenetten (transport, distributie, aansluit) in de bebouwde omgeving beïnvloeden. Deze is opgesteld op basis van beschikbare literatuur, aangevuld met de uitkomsten (data analyse en praktijk kennis) van dit project.
- ✓ **Een kostenmodel** wat op basis van de bovengenoemde database kosteninschattingen genereert voor de aanleg van warmtenetten in de bebouwde omgeving. Bij het opzetten van het model wordt rekening gehouden met een variatie aan omgevingsfactoren. Specifieke verschillen tussen HT, MT, en LT, warmtenetten worden niet onderzocht omdat de beschikbare data hiervoor niet toereikend was.
- ✓ **Een rapportage** van de resultaten.

## 2 Samenvatting literatuurstudie

Het volgende hoofdstuk geeft een overzicht van de onderzochte literatuur en bestaande kennis op verschillende sub onderwerpen die voor deze studie relevante zijn. Dit hoofdstuk is een belangrijke basis voor de (methodische) opzet van het onderzoek alsook voor de reeds bekende en onderzochte omgevingsvariabelen. Het is geen deliverable voor project 2C. Echter, vanwege de relevantie voor de methodische opbouw en uitgesproken interesse door verschillende stakeholders wordt een samenvatting van de gedane literatuurstudie beschikbaar gesteld in deze rapportage.

Voor meer informatie zie de afstudeer scriptie van Jorn Mieras waar een gedetailleerde literatuurstudie wordt gepresenteerd (Mieras, 2021) of de artikelen waar naar verwezen wordt voor de verschillende sub-onderwerpen.

### 2.1 Technieken voor kostenramingen

In de literatuur worden verschillende technieken genoemd en vergeleken om kostenramingen op te stellen. De meeste publicaties komen van oorsprong uit het ruimtevaart domein. De drie meest voorkomende, ofwel klassieke technieken voor kostenraming zijn analoog, parametrisch en bottom-up (Curran et al., 2004). Daarnaast is er nog een veel gebruikte categorie, meer van kwalitatieve aard, die uitgaat van 'expert judgement' of 'intuïtie'. Het gebruik van rekenregels of kentallen valt hier bijvoorbeeld ook onder. Elke aanpak heeft zijn voor en nadelen en onderliggende aannames (Duran et al., 2008; Hueber et al., 2016).

- i. *Analoog. Een vergelijkbare dienst of product wordt gebruikt als de basis voor kostenramingen en deze wordt aangepast aan de hand van de verschillen tussen beide.*
- ii. *Parametrisch. Tussen het product/dienst en de bijbehorende eigenschappen worden relaties bepaald op basis van historische kosten en eigenschappen. Deze relaties worden gebruikt voor het schatten voor de kostenramingen.*
- iii. *Bottom-up. Een optel som van alle losse kostencomponenten (zoals arbeid, materiaal, etc.) die in het productie proces worden gebruikt leidend tot een kostenraming.*

Het huidige onderzoek maakt vooral gebruik van de parametrische aanpak waarin de relatie tussen omgevingseigenschappen en aanlegkosten centraal staat. In de literatuurstudie wordt ook geïnventariseerd welke factoren van invloed zijn op de kosten van water en gas infrastructuur om op een analoge wijze een startpunt te creëren voor het definiëren van de databehoeft (zie 3.1). Parametrisch modelleren heeft als voordeel dat deze techniek geschikt is in de vroege plan en project fases, wanneer een gedetailleerd technisch ontwerp nodig voor bijvoorbeeld een bottom-up aanpak, nog niet beschikbaar is (Dai, 2016). Naar verwachting zullen de resultaten en opgedane inzichten van deze aanpak ook input zijn voor een bottom-up kostenraming.

## 2.2 Statistische modellen voor kostenramingen van warmtenetten

Er zijn weinig wetenschappelijke studies gedaan waarin specifiek omgevingsfactoren meegenomen worden in het opstellen van kost functies voor warmtenetten. De enige studie waarin dit wel is gebeurd toont aan dat de aanwezigheid van teer in het asfalt van significante invloed was op de aanlegkosten voor wijken met vrijstaande woningen (Reidhav and Werner, 2006). Onderzochte variabelen waren aansluit ratio (aangesloten huizen/ totaal aantal huizen in het gebied), leiding lengte per huis, vorst aan de grond, teer in asfalt, stenen aanwezig, en over dimensionering in het verleden. Dit is gebaseerd op een database van 55 projecten in Göteborg in Zweden en levert het volgende model voor de aanlegkosten (€/woning) met een R2 waarde van 0.7.

$$Ch_{sweden} = 4230 + l_{house} * 232 + boolean_{asphalt} * 2360$$

Enkel de aanwezigheid van asfalt, en leiding lengte per huis bleken significant te zijn in de correlatie met de kosten. Overige variabelen kunnen ook nog steeds relevant zijn, bij meer of andere projecten in de data set bijvoorbeeld.

Multivariate regressie analyse is de meest gebruikte aanpak, voor warmtenetten en andere infrastructuur (water en gas), om kosten te voorspellen. Bij kleine en lineaire datasets werkt deze aanpak relatief beter dan andere modellen. Daarnaast kan dit model het makkelijkst worden geïnterpreteerd. Aandachtspunten bij deze aanpak zijn i) alleen significante variabelen opnemen in model, te veel variabelen verslechtert de prestatie ii) interactie termen<sup>1</sup> lijken belangrijk.

Er komen veel relatief kleine datasets voor in de wetenschappelijk gepubliceerde onderzoeken (55, 17, 50, en 130 projecten) voor waternet, en warmtenet onderzoeken. De conclusies zijn vaak voorzichtig geformuleerd, met sterke aanbeveling voor meer data voor sterkere conclusies.

Er wordt geconstateerd, in een onderzoek in Zweden, dat het onderzoeken en vergelijken van kosten van verschillende bedrijven moeilijk is. Verschillende bedrijven nemen andere kosten componenten mee in hun "aanleg kosten". Hier speelt dus ook het probleem dat het moeilijk is om 'appels met appels' te vergelijken wanneer er data wordt geanalyseerd van verschillende organisaties.

## 2.3 Omgevingsfactoren in energie transitie modellen

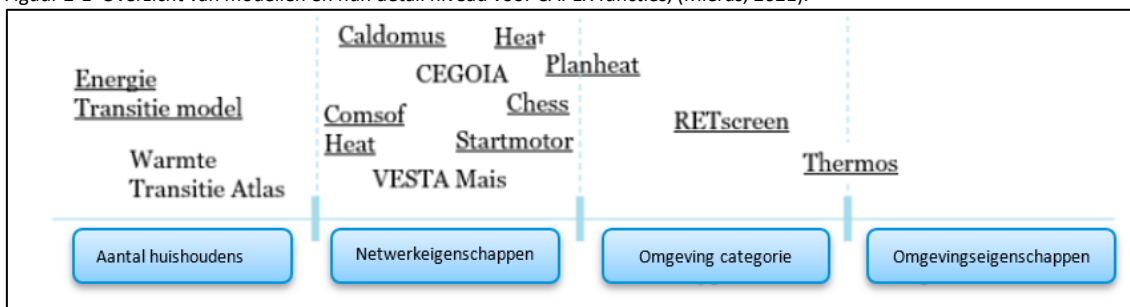
Er zijn veel verschillende energie transitie modellen die ontwikkeld zijn om beleidsmakers te ondersteunen in de keuzes voor transitiepaden en de ontwikkeling van energie transitie strategieën. De modellen variëren van elkaar, hoofdzakelijk op het detail niveau van (kosten)berekeningen, het kunnen meenemen van toekomstige scenario's en de hoeveelheid en variatie aan energiebronnen die onderzocht en vergeleken kunnen worden.

Aansluitend bij het doel van dit onderzoek worden enkel modellen beschouwd die CAPEX kostenramingen maken in een vroege project fase en waarvoor beschrijvingen of toelichtingen van de werking van het model, inclusief de aannames, openbaar beschikbaar en toegankelijk zijn. Wanneer er wordt ingezoomd op hoe de verschillende modellen de CAPEX kosten berekenen is te

<sup>1</sup> Interaction terms: creëren van niet lineaire input variabelen door twee onafhankelijke input variabelen met elkaar te vermenigvuldigen. Bijvoorbeeld, de kosten voor het graven van een sleuf worden bepaald door de diepte en de diameter tezamen.

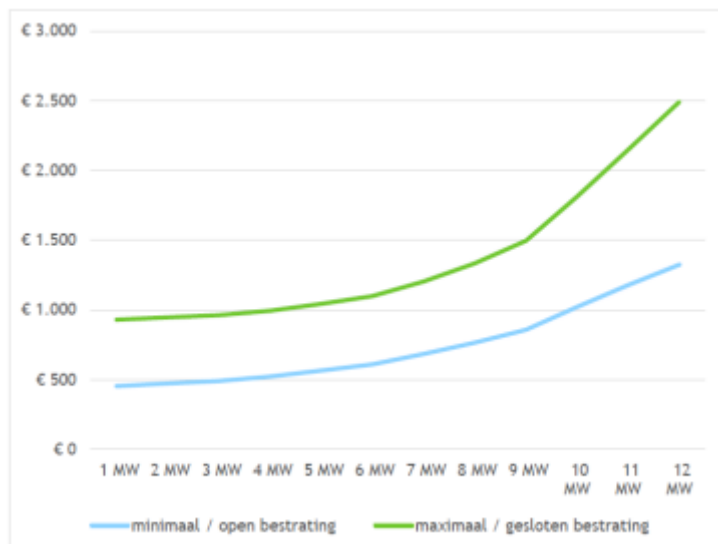
zien dat het merendeel van de modellen de CAPEX kosten baseert op technische eigenschappen van het netwerk, aanvullend op het aantal aan te sluiten huishoudens (Figuur 2-1 ). In sommige gevallen worden de gebruikte kentallen wel aangepast naar het type omgeving waarin het betreffende netwerk aangelegd moet worden door bijvoorbeeld een hogere meter prijs te gebruiken in stedelijk gebied ten opzichte van landelijk gebied. Er is geen model gevonden waarin specifieke omgevingseigenschappen zoals het type bestrating, beschikbare ruimte, etc, gebruikt worden als model input.

Figuur 2-1 Overzicht van modellen en hun detail niveau voor CAPEX functies, (Mieras, 2021).



In verschillende omschrijvingen wordt wel gesuggereerd dat een deel van de onzekerheid (bandbreedte) van de voorspelde kosten te maken hebben met omgevingsfactoren. In de documentatie van VESTA, meest gebruikt in de Nederlandse energie transitie context, wordt ook aangekaart dat voorspellingen te maken hebben met een grootte bandbreedte van onzekerheid, vooral bepaald door het type bestrating (Figuur 2-2). De bovengrens is ongeveer twee keer zo hoog als de ondergrens, afhankelijk van het vermogen. Ook worden factoren genoemd zoals drukte in de ondergrond, en grondvervuiling, die veel invloed hebben op variaties in de kosten.

Figuur 2-2 Vesta model – Boven en ondergrens kosteninschattingen



Tabel geeft een overzicht van de verschillende kostenfuncties die gebruikt worden in verschillende modellen.

Tabel 2.1 CAPEX kostenfuncties voor Energie Transitie Modellen

Model	CAPEX functie																				
Energie Transitie Model	$CAPEX = CAPEX_{transport} + CAPEX_{distribution} + CAPEX_{stations}$ $CAPEX_{transport} = C_{p_{transport}} * P_{peak}$ $CAPEX_{distribution} = C_{m_{distribution}} * \overline{L_{connection}} * N_{households}$ $CAPEX_{stations} = (C_{p_{HTS}} + C_{p_{sub}}) * P_{peak}$																				
Warmte Transitie Atlas	$CAPEX = Ch_{rural/urban} * N_{households}$ $Ch_{rural/urban} = f(B_{type}, B_{year})$																				
Caldomus	$CAPEX = CAPEX_{transport} + CAPEX_{distribution} + CAPEX_{stations}$ $CAPEX_{transport} = L_{trans\_network} * (\Delta T)^{-0.189} * 624.62 * (P_{peak})^{0.1893}$ $CAPEX_{distribution} = C_{m_{distribution}} * L_{dis\_network}$ $L_{dis\_network} \subset \begin{cases} primary_{dis} = f_{pri} * 2 * \sqrt{2} * \sqrt{A_{network}} \\ secondary_{dis} = f_{sec} * N_{sub\_station} * 0.25 * 0.5 * \sqrt{2} * \sqrt{A_{network}} \end{cases}$ $CAPEX_{stations} = (C_{p_{HTS}} + C_{p_{sub}}) * P_{peak}$																				
VESTA	$CAPEX = Cost_{man} + Cost_{com} + CAPEX_{transport} + CAPEX_{distribution}$ $+ \sum_{connection=1}^{N_{households}} CAPEX_{connection} + CAPEX_{stations} + CAPEX_{buffer}$ $CAPEX_{transport} = C_{m_{pipe}}(P_{peak}) * L_{source} * f_{detour}$ $CAPEX_{distribution} \subset \begin{cases} WKO\&TEO = (600 * \frac{187}{15741}) * A_{network} \\ LT = C_{m_{pipe}}(P_{peak}) * f_{LT} * \sqrt{A_{network}} \\ MT = C_{m_{pipe}}(P_{peak}) * f_{MT} * L_{road} \end{cases}$ $CAPEX_{connection} = L_{dis\_road} * C_{m_{pipe}}(P_{connection}) * N_{connections}$ $CAPEX_{stations} = (C_{p_{HTS}} + C_{p_{sub}}) * P_{peak}$ $CAPEX_{buffer} = Ch_{buffer} * N_{households}$ $C_{m_{pipe}}(P) \subset \begin{cases} min = 400 + 210 * (P * 0.001)^{0.5} \\ max = 800 + 200 * (P * 0.001)^{0.6} \end{cases}$																				
RETscreen	$CAPEX = \sum_{i=1}^{N_{trans\_seg}} CAPEX_{transport} + \sum_{j=1}^{N_{dis\_seg}} CAPEX_{distribution} + \sum_{k=1}^{N_{stations}} CAPEX_{stations}$ $CAPEX_{transport} = f(L_i, S_{type}, d_i, f_{trans})$ $CAPEX_{distribution} = f(L_j, S_{type}, d_j, f_{dis})$ $CAPEX_{stations} = f(S_{type}, f_{stations}, N_{stations})$																				
Thermos	$CAPEX = \sum_{i=1}^{N_{pipe\_segments}} Mechanical_{cost}(i) + \sum_{i=1}^{N_{pipe\_segments}} Civil_{cost}(i)$ $Mechanical_{cost}(i) = L_i * (f_{mech\_1}(i) + (f_{mech\_2}(i) * d_i)^{1.3})$ $Civil_{cost}(i) = L_i * (f_{civ\_1}(i) + (f_{civ\_2}(i) * d_i)^{1.1})$ <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>Location</th> <th>Surface</th> <th><math>f_{civ\_1}</math></th> <th><math>f_{civ\_2}</math></th> </tr> </thead> <tbody> <tr> <td>Urban</td> <td>Hard</td> <td>1200</td> <td>500</td> </tr> <tr> <td>Urban</td> <td>Soft</td> <td>450</td> <td>0</td> </tr> <tr> <td>Suburban</td> <td>Hard</td> <td>850</td> <td>200</td> </tr> <tr> <td>Suburban</td> <td>Soft</td> <td>100</td> <td>0</td> </tr> </tbody> </table> <p>Sensible civil parameters for the UK:</p>	Location	Surface	$f_{civ\_1}$	$f_{civ\_2}$	Urban	Hard	1200	500	Urban	Soft	450	0	Suburban	Hard	850	200	Suburban	Soft	100	0
Location	Surface	$f_{civ\_1}$	$f_{civ\_2}$																		
Urban	Hard	1200	500																		
Urban	Soft	450	0																		
Suburban	Hard	850	200																		
Suburban	Soft	100	0																		



## 2.4 Kostendrijvers

De factoren van invloed die in de literatuurstudie naar voren zijn gekomen en/of reeds gebruikt worden in de energie transitie modellen:

Figuur 2-3 Overzicht van de verschillende factoren die in verschillende literatuurstudies zijn onderzocht of in rekenmodellen zijn opgenomen. Zie (Mieras, 2021) voor een volledige beschrijving.

Energie transitie modellen	Cost estimation warmtenetten	Cost estimation vergelijkbare infra
<ul style="list-style-type: none"> <li>➤ Aantal huishoudens aangekoppeld</li> <li>➤ temperatuur van het warmtenet (warmtevraag)</li> <li>➤ lengte van de leidingen</li> <li>➤ diameter van de leidingen</li> <li>➤ Stedelijk of landelijke omgeving</li> <li>➤ Wegbekleding</li> <li>➤ Verkeersdrukke</li> <li>➤ kwaliteit van de grond</li> <li>➤ type gebouwen</li> <li>➤ drukke in de ondergrond</li> <li>➤ Grond vervuiling</li> </ul>	<ul style="list-style-type: none"> <li>➤ aanwezigheid van teer in het asfalt</li> <li>➤ aansluit rato</li> <li>➤ leiding lengte per huis</li> <li>➤ vorst aan de grond</li> <li>➤ aanwezigheid van stenen</li> <li>➤ overdimensionering (voor toekomstig gebruik)</li> <li>➤ diameter van de leiding</li> <li>➤ zandvulling</li> <li>➤ leidingen met 200mm isolatie</li> </ul>	<ul style="list-style-type: none"> <li>➤ diepte uitgraving</li> <li>➤ verschillende materiaaltypen hebben verschillende kostfuncties</li> <li>➤ aanleg van nieuwe leidingen</li> <li>➤ aanleg van zijleiding</li> <li>➤ verlaten leidingen</li> <li>➤ punt reparaties</li> <li>➤ mangaten</li> <li>➤ voetgangershellingsen</li> <li>➤ trottoirs</li> <li>➤ stoepranden</li> <li>➤ asfalt en beton verhardingen</li> <li>➤ asfalt overlay</li> <li>➤ resurfacing' materialen</li> <li>➤ slurry afdichting</li> <li>➤ asfalt patchen</li> <li>➤ uitgegraven grond</li> <li>➤ geïmporteerde grond</li> <li>➤ beddingmateriaal</li> <li>➤ brandkraanpunten</li> <li>➤ schuifafsluiters</li> <li>➤ uitblaasleiding</li> </ul>

## 3 Opstellen database

### Opstellen database

- Definiëren databehoefte en databeschikbaarheid
- Ontwikkelen template en GIS script
- Data collectie
- Opschonen database

Deel een van het project betrof het opstellen van een database met gegevens over historische (in het verleden uitgevoerde en afgeronde) warmtenet aanlegprojecten, in de bebouwde omgeving.

### 3.1 Definiëren databehoefte en databeschikbaarheid

Het startpunt voor het definiëren van de data behoefte is het resultaat van de literatuurstudie. Daarin zijn de reeds bekende factoren die invloed hebben op de aanlegkosten van leidinginfrastructuur geïdentificeerd. Vervolgens zijn er aanvullingen gedaan op basis van de praktijkervaring van de projectpartners. Bijvoorbeeld, de aanwezigheid van “Niet Gesprongen Explosieven” in een projectgebied leidt volgens de projectpartners tot verhoging van de aanlegkosten door aanvullende maatregelen en mogelijk oponthoud. Andere voorbeelden zijn het moeten kruisen van andere aanwezige infrastructuur (zoals riolering of tramrails), toepassen van bemaling, of het toepassen van grondverbetering als gevolg van vervuilde grond. Het samenvoegen van al deze variabelen heeft geleid tot een “longlist” van factoren, opgenomen in Tabel 3.1, die volgens de literatuur, aangevuld met praktijk kennis van de partners, van invloed (kunnen) zijn op de aanlegkosten en daarmee, interessant zijn om op te nemen in de database om statistische verbanden nader te kunnen onderzoeken.

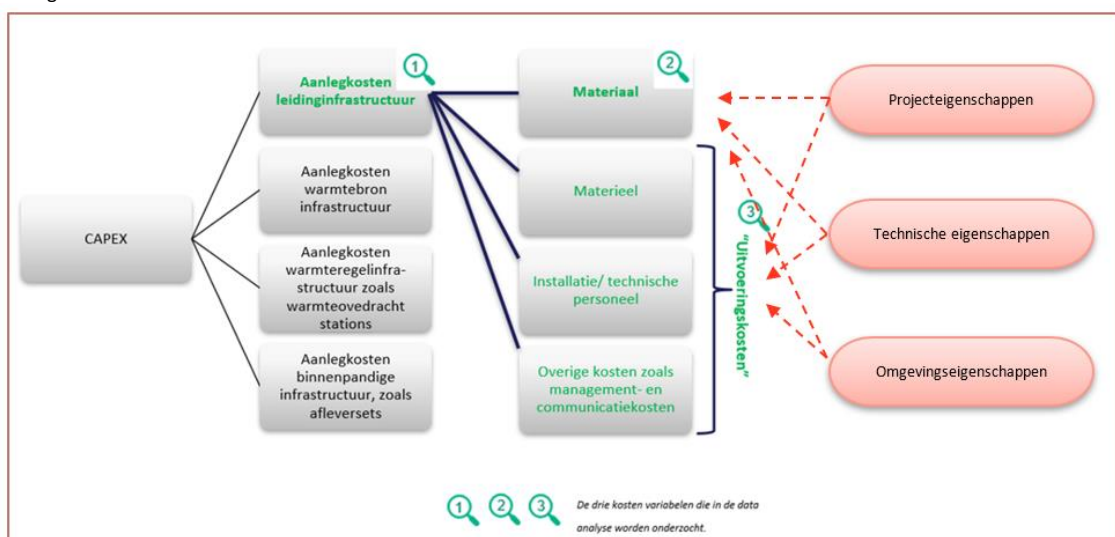
Deze “longlist” van factoren kan worden gezien als een soort wensenlijst. In een ideale situatie wordt voor elke factor een variabel gevonden die statistisch onderzocht wordt op de invloed op de aanlegkosten. Echter, deze “longlist” betreft factoren waarvoor dus nog een vertaalslag naar meetbare variabelen moet worden gemaakt, en daarmee is dit niet de lijst van variabelen die daadwerkelijk onderzocht wordt in de statistische analyse. Dit komt onder andere doordat niet voor alle factoren data beschikbaar is in een bruikbaar format. In sommige gevallen, zoals bijvoorbeeld voor “drukke in de ondergrond” is geen uniforme variabel beschikbaar die dit weergeeft. Daarom is er gekeken naar variabelen die als proxy kunnen dienen. Dergelijke aanpassingen en overwegingen hebben geleid tot een “longlist” aan variabelen, te zien in bijlage A, die verzameld worden in de data uitvraag. In deze bijlage wordt ook een specifiekere bronverwijzing gegeven naar de diverse GIS databases die gebruikt zijn en een toelichting op verschillende type variabelen. Deze “longlist” aan variabelen wordt na de data verzameling opgeschoond om te komen tot een “shortlist” aan variabelen Tabel 3.2. Nadere toelichting over de opschoning is te lezen in 3.4.

De verschillende variabelen zijn voor dit onderzoek, ten behoeve van de overzichtelijkheid, onderverdeeld in verschillende categorieën namelijk projecteigenschappen, technische eigenschappen en omgevingseigenschappen. Daarnaast is er nog de categorie aanlegkosten, waar

de kosten posten “Aanlegkosten leidinginfrastructuur”, “Materiaal” en “Uitvoeringskosten” onder vallen (zie ook 1.3). Hiermee wordt een database opgebouwd waarmee in de tweede fase van dit project onderzocht kan worden hoe elke van de variabelen in deze categorieën een statistische relatie hebben met de kosten posten, zoals te zien in Figuur 3-1. De nadruk van dit onderzoek ligt bij het identificeren van de invloed van de omgevingseigenschappen op de aanlegkosten. Echter, projecteigenschappen en technische eigenschappen zullen uiteraard ook van invloed zijn op de aanlegkosten (mogelijk in samenhang met omgevingseigenschappen) en moeten daarom ook in de modelanalyse meegenomen worden.

Er worden twee hoofdbronnen geïdentificeerd voor de verzameling van de data en het opstellen van de database: het verzamelen van gegevens bij de warmtebedrijven en het verzamelen van gegevens uit nationale GIS databases

Figuur 3-1 Overzicht en relatie categorieën projecteigenschappen, technische eigenschappen, omgevingseigenschappen, en aanlegkosten.



Tabel 3.1 Longlist van factoren met invloed op de CAPEXfuncties. PE = Project eigenschappen, TE = technische eigenschappen, OE = omgevingseigenschappen. Literatuur = Factor van invloed geïdentificeerd in literatuur, Praktijk = factor van invloed aanvullend geïdentificeerd door praktijk kennis.








Categorie	Factor van invloed	Literatuur	Praktijk
PE	Aanleg nieuwe leidingen of vervangen oude leidingen	✓	
PE	Aanwezigheid van verlaten leidingen	✓	
PE	Datum aanleg		✓
PE	Duur aanleg		✓
PE	Type contract (bijv. aanleg en onderhoud)		✓
PE	Andere werkzaamheden, bijvoorbeeld aan het riool, 'mee gekoppeld'		✓
TE	Aantal huishoudens / aansluitingen	✓	
TE	Warmtevraag / temperatuur	✓	
TE	Leiding lengte totaal	✓	
TE	Leiding lengte per huis	✓	
TE	Leiding type (aansluit, distributie, transport)	✓	
TE	Leiding diameter	✓	
TE	Aansluit rato	✓	
TE	Diepte leiding ligging	✓	
TE	Overdimensionering (voor toekomstig gebruik)	✓	
TE	Gebruikt leiding (en/of isolatie) materiaal	✓	



TE	Aanleg van zijleidingen	✓	
TE	Gebruikte aanleg methode		✓
TE	Leiding lengte door voortuin (aansluit leidingen)		✓
OE	Trottoirs	✓	
OE	Voetgangershellingen aanwezig	✓	
OE	stoepranden	✓	
OE	Aanwezigheid van stenen (in de bodem)	✓	
OE	Vorst aan de grond	✓	
OE	Type gebouw	✓	
OE	Drukke in de ondergrond	✓	
OE	Kruising met andere infrastructuur	✓	
OE	Verkeersdrukke	✓	
OE	Wegbekleding (open / gesloten)	✓	
OE	Wegbekleding (asfalt met teer)	✓	
OE	Stedelijk / Landelijke omgeving	✓	
OE	Uitgraven grond	✓	
OE	Geïmporteerde grond / grondverbetering	✓	
OE	Grondvervuiling / kwaliteit	✓	
OE	Postcode		✓
OE	Type weg (bijvoorbeeld, straat of provinciale weg)		✓
OE	Bomen		✓
OE	Niet gesprongen explosieven		✓
OE	Archeologie		✓
OE	Bemaling toegepast		✓
OE	Grondwaterstand		✓
OE	Herbestrating uitgevoerd (door dezelfde partij(grondroerder)		✓
OE	Kruising met waterlichaam of brug		✓
OE	Waterlichaam in nabijheid		✓

Tabel 3.2 "Shortlist" van variabelen. PE = Project eigenschappen, TE = technische eigenschappen, OE = omgevingseigenschappen, Bron WB = Warmtebedrijf, Bron GIS = Geografisch informatie systeem, ● = Nominaal, ■ = Ordinaal, 📏 = Schaal

Categorie	Variabel naam	Variabel omschrijving en eenheid	Bron	Type
PE	Bedrijf	Het warmtebedrijf dat het project gerealiseerd heeft en de data heeft aangeleverd. Geanonimiseerd	WB	●●●
PE	Type	Soort aanleg project. 1 = Vervanging bestaand warmtenetwerk, 2 = Aanleg inbreidingslocatie, 3 = Aansluiting bestaande bouw	WB	●●●
PE	Contract	Welke werkzaamheden waren gecontracteerd? 1 = Uitvoering, 2 = Engineering en uitvoering, 3 = Ontwerp, Engineering en uitvoering	WB	●●●
PE	Geplande duur	Hoe lang het aanlegproject naar verwachting zou gaan duren, in dagen	WB	📏
PE	Tijdstempel	Berekend op basis van de startdatum van het project, geeft de startdatum aan op een doorlopende tijdschaal. Het meest recente project uit de database is gestart in aug 2020 en het oudste project is gestart in januari 2015	WB	📏
TE	Supply temp max	De hoogst opgegeven supply temperatuur (deze was opgevraagd voor verschillende delen van het netwerk) in graden celsius	WB	📏
TE	Delta temp max	Her verschil tussen supply en retour temperatuur, hoogste waarde van de verschillende delen van het netwerk, in graden celsius	WB	📏
TE	Lengte	Totale lengte van de aangelegde leidingen	WB	📏

TE	Aantal afnemers	Het aantal afnemers dat is aangesloten in dit project. (bij een flatgebouw gaat het dus om het aantal aangesloten appartementen) .	WB	
TE	Aantal aansluitingen	Het aantal aansluitleidingen naar het distributienet (bij een flatgebouw is dit dus 1 aansluiting)	WB	
TE	Aangesloten vermogen	Max vermogen dat is aangesloten op het warmtenet – geeft samen met het max aangelegd vermogen een beeld van de toekomstbestendigheid (kW)	WB	
TE	Aangelegd vermogen	Max vermogen waar het warmtenet op gedimensioneerd is. Als dit hetzelfde is als het aangesloten vermogen kunnen er geen nieuwe aansluitingen meer bij op het net. (kW)	WB	
TE	Ruimte op het net	Berekend - verschil tussen Aangelegd en Aangesloten vermogen. 0 = nee (aangelegd vermogen volledig aangesloten), 1 = ja (aangelegd vermogen nog niet volledig aangesloten)	WB	
TE	Dominant leiding type	Meest voorkomende leidingtype in de projectzone. Distributieleiding=1, Aansluitleiding = 2, Transportleiding = 3	WB	
TE	Dominant leiding materiaal	Meest voorkomende leidingmateriaal in de projectzone. Staal-PUR-PE-VPS=1, Steelflex -flexibel staal=2, Calpex PEX_PUR_PE = 3, Overig = 4.	WB	
TE	Leiding diameter	Gemiddelde diameter van de leidingen in de projectzone, in millimeters (mm)	WB	
TE	Methode Opensleuf	Opensleuf is gebruikt als aanlegmethode voor de leidingen in de projectzone 0 = nee, 1 = ja	WB	
TE	Diepte Opensleuf	De diepte van de sleuf, in meters (m).	WB	
TE	Methode: Bovengronds	Er is bovengronds gewerkt voor het aanleggen van de leidingen in de projectzone 0 = nee, 1 = ja	WB	
OE	Postcode	Postcode gebied (2 cijfers) waar het project is uitgevoerd.	GIS	
OE	Grondvervuiling	De intensiteit van de grondvervuiling die aanwezig is het projectgebied. geen=0, licht=1, middel=2, zwaar=3. Categorieën bekend en gebruikt bij warmtebedrijven	WB	
OE	Lengte van vervuild tracé	% van het projectgebied dat te maken had met vervuild oppervlak. Berekend op basis van vervuilingsoppervlak en totale tracélengte	WB	
OE	Aandeel vervuilde grond	Fractie vervuilde grond van totaal leiding lengte, tussen 0 en 1	WB	
OE	Grondverbetering	Is er tijdens de uitvoeringswerkzaamheden grondverbetering toegepast? 0 = nee, 1 = ja	WB	
OE	Grondwaterniveau	De hoogte van het grondwater in de projectzone, in meters (m)	GIS	
OE	Bemaling	Is er tijdens de uitvoeringswerkzaamheden bemaling toegepast? 0 = nee, 1 = ja	WB	
OE	Straat	Het wegtype 'straat' is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Weg van zeer plaatselijk belang, gelegen binnen bebouwd gebied</i>	GIS	
OE	Hoofdweg	Het wegtype 'hoofdweg' is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Verharde weg die is aangeduid met een E-nummer, maar niet met een A-nummer, of verharde weg die onderdeel is van een verbindingsroute tussen grotere plaatsen, wat blijkt uit blauwe ANWB-borden, dan wel onderdeel is van een route om eindigende A of E-routes tot een gesloten netwerk te completeren</i>	GIS	
OE	Regionale weg	Het wegtype 'regionale weg' is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Verharde weg die een verbinding vormt tussen bewoonde oorden of grote stadswijken, en daartoe van twee kanten bewegwijzerd zijn met blauwe ANWB-richtingsborden voor autoverkeer</i>	GIS	

OE	Lokale weg	Het wegtype 'lokale weg' is aanwezig in de projectzone, 0 = nee, 1 = ja, <i>Weg van lokaal belang tussen bewegwijzerde routes</i>	GIS	
OE	Wegtype verbinding voetafdruk	Fractie van het projectgebied met het wegtype 'verbinding' (tussen 0 en 1) <i>Elk wegdeel wordt voorzien van een label waarbij de keuze bestaat uit 'verbinding', 'kruising' en 'overig verkeersgebied'</i>	GIS	
OE	Wegtype kruising voetafdruk	Fractie van het projectgebied met het wegtype 'kruising' (tussen 0 en 1) <i>Elk wegdeel wordt voorzien van een label waarbij de keuze bestaat uit 'verbinding', 'kruising' en 'overig verkeersgebied'</i>	GIS	
OE	Wegtype overig verkeersgebied voetafdruk	Fractie van het projectgebied met het wegtype 'overig verkeersgebied' (tussen 0 en 1) <i>Elk wegdeel wordt voorzien van een label waarbij de keuze bestaat uit 'verbinding', 'kruising' en 'overig verkeersgebied'</i>	GIS	
OE	Wegbreedte	De gemiddelde breedte van de weg in het projectgebied, in meters (m) <i>Weg is gebaad gedeelte voor het wegverkeer en kan meerdere types bevatten (zie wegtype)</i>	GIS	
OE	Wegbreedte: >7m	Een weg breder dan 7m is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Weg is gebaad gedeelte voor het wegverkeer en kan meerdere types bevatten (zie wegtype)</i>	GIS	
OE	Wegbreedte: 4-7m	Een weg met een breedte tussen de 4 en 7m is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Weg is gebaad gedeelte voor het wegverkeer en kan meerdere types bevatten (zie wegtype)</i>	GIS	
OE	Wegbreedte: 2-4m	Een weg met een breedte tussen de 2 en 4m is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Weg is gebaad gedeelte voor het wegverkeer en kan meerdere types bevatten (zie wegtype)</i>	GIS	
OE	Wegbreedte: <2m	Een weg smaller dan 2m is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Weg is gebaad gedeelte voor het wegverkeer en kan meerdere types bevatten (zie wegtype)</i>	GIS	
OE	Weg voetafdruk	Fractie van de projectzone met weg (tussen 0 en 1). <i>Weg is gebaad gedeelte voor het wegverkeer en kan meerdere types bevatten (zie wegtype)</i>	GIS	
OE	Gebouw voetafdruk	Fractie van de projectzone met gebouw (tussen 0 en 1) <i>Gebouw is Pand of Verblijfsobject</i>	GIS	
OE	Bouwjaar	Het gemiddelde bouwjaar van de gebouwen in de projectzone	GIS	
OE	Standaard deviatie Bouwjaar	De standaard deviatie van het gemiddelde bouwjaar van de gebouwen in de projectzone (uitgedrukt in aantal jaren)	GIS	
OE	Afstand tussen gebouw en weg	De gemiddelde afstand tussen gebouw en weg in de projectzone in meters (m)	GIS	
OE	Inwonersdichtheid	Gemiddelde inwonersdichtheid in de projectzone (#/hectare)	GIS	
OE	Woongebied	Projectzone valt onder het type 'woongebied', 0 = nee, 1 = ja	GIS	
OE	Spoorlijn	Een spoorlijn is aanwezig in de projectzone, 0 = nee, 1 = ja <i>Gebaand gedeelte voor het verkeer over rails (Voorbeelden zijn: spoorbaan, metrobaan, trambaan)</i>	GIS	
OE	Waterlichaam	Een oppervlaktewaterlichaam (sloot, meer, rivier,...) is aanwezig in de projectzone, 0 = nee, 1 = ja	GIS	
OE	Peilgebied	Het peil van het peilgebied waar de projectzone in valt.	GIS	

OE	Voetafdruk bomen	% van de project zone bedekt met bomen (waarbij het gaat om bomen >2,5 m)	GIS	
OE	Trefkans archeologie	de trefkans (hoog, middelhoog, laag) op archeologische resten.	GIS	

## 3.2 Ontwikkelen template

Voor het verzamelen van data van de verschillende warmtebedrijven is gezamenlijk een template opgesteld op basis waarvan de warmtebedrijven de projectgegevens op een uniforme wijze aan konden leveren. Het was zeer van belang dat de invulvelden op de juiste manier zouden worden geïnterpreteerd en dat de data van het ene bedrijf dezelfde inhoud reflecteert als de data van het andere bedrijf. Door de projectpartners is meer dan eens de zorg uitgesproken over de vergelijkbaarheid van de data: *“we willen appels met appels vergelijken en niet appels met peren”*. Ook uit de literatuur kwam naar voren dat het vergelijken van kosten van verschillende bedrijven moeilijk is omdat kostenposten en definities onderling variëren.

Er is veel aandacht besteed in het project aan de ontwikkeling van de template en het beheersen van het hierboven beschreven risico. Enerzijds door de template in nauwe samenwerking met de warmte bedrijven te ontwikkelen en testen – *hierdoor zijn ook de ‘invullers’ op de hoogte van de opbouw en de intentie van de template en de betekenis van de invulvelden* – en anderzijds door de template zo helder mogelijk op te zetten – *genoeg informatie voor de juiste interpretatie maar ook geen informatie overload*.

Er is gebruik gemaakt van een aantal pilot projecten om de bruikbaarheid en nauwkeurigheid van de template te toetsen voor de verdere, grootschalige, uitvoering. Het invullen van de template door de warmtebedrijven kostte ongeveer 1 a 2 uur per project. In sommige gevallen moesten ook verschillende interne informatie systemen geraadpleegd worden om de nodige informatie te vinden.

Figuur 3-2 geeft een impressie van de opzet van de template. De overige onderdelen van de template zijn te vinden in Bijlage B. Er is een script (Python) ontwikkeld om alle informatie uit de ingevulde templates te extraheren en samen te voegen tot een bestand. De template en het script kunnen worden opgevraagd bij Deltares.

Figuur 3-2 Voorbeeld template onderdeel A – projecteigenschappen

**WARMINGUP**  
Deltares

Gegevens hoofdproject

#N/A

Project	#N/A
Project ID	#N/A
Detailnummer	#N/A

Onderdeel A  
project eigenschappen

**A.1-Locatie**

Latitude [deg]	Longitude [deg]	
		Vergeet niet om een afbeelding van het tracé (GIS / technische tekening o.i.d.) mee te sturen met dit excell!

**A.2-Planning**

	Startdatum	Einddatum
Werkelijk		
Gepland		

**A.3-Grondvervuiling**

*Geef aan op hoeveel meter van het tracé grondvervuiling wel of niet een rol heeft gespeeld.*

Categorie vervuiling	Meter tracé [m]
Onbekend	
0m	

Totale tracé lengte:

**A.4-Niet gesprongen explosieven (NGE)**

Zijn er tijdens de uitvoering additionele en/of voorzorgsmaatregelen getroffen omdat er mogelijk NGE aanwezig waren?

**A.5- Archeologie**

Zijn er tijdens de uitvoering additionele en/of voorzorgsmaatregelen getroffen omdat er archeologische vondsten gedaan zijn?

**A.6-Meekoppelen andere infra**

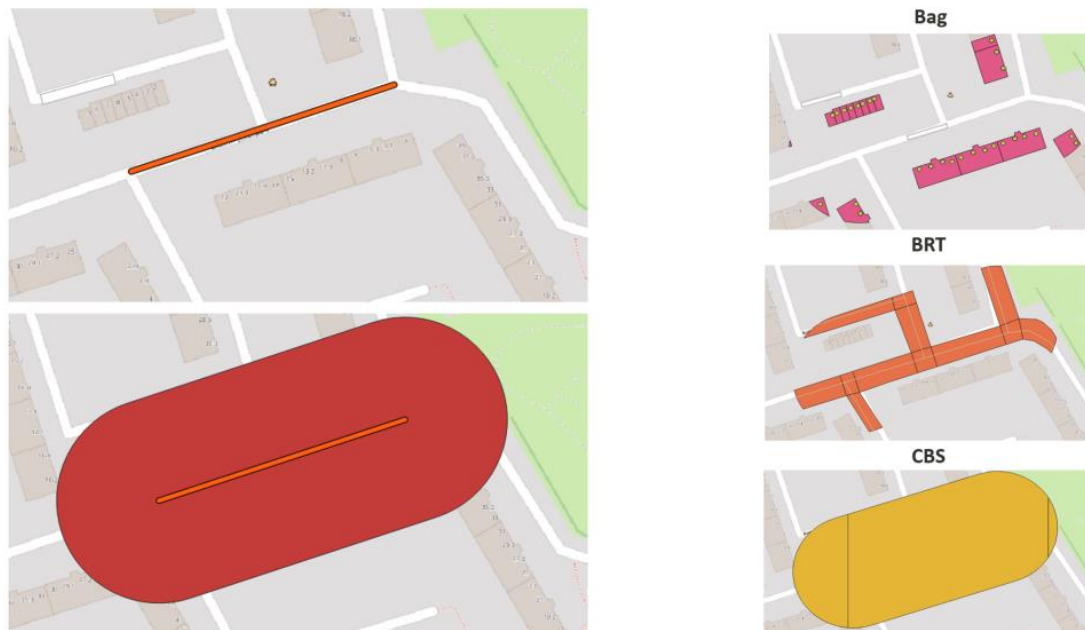
Zijn er tijdens de aanleg van het warmtenet gelijktijdig andere werkzaamheden uitgevoerd (bijvoorbeeld het vervangen van andere infra)?

### 3.3 GIS data

Naast het verzamelen van de nodige gegevens bij de warmtebedrijven zijn ook gegevens over de projectlocaties uit Nationale GIS databases verzameld. Dit proces is weergegeven in Figuur 3-3. Op basis van de project locatiegegevens (ontvangen van de warmtebedrijven) zijn de betreffende leiding tracés ingetekend en opgeslagen als shapefiles. Rondom het ingetekende tracé heen wordt een projectzone gecreëerd. Deze projectzone wordt vervolgens gebruikt om voor het projectgebied de gegevens uit de nationale databases te importeren. Deze projectzone beslaat een oppervlak van 30meter rondom het tracé. Er zijn verschillende groottes onderzocht – te klein betekent dat locatie gegevens mogelijk niet ‘gevangen’ worden, en te groot betekent dat er te veel gegevens ‘gevangen’ worden binnen de project zone. Er is een gevoeligheidsanalyse uitgevoerd voor de te selecteren grootte van de projectzone (bijlage C).



Figuur 3-3 Importeren GIS data voor project tracés. Linksboven: tracé, linksonder: projectzone, rechts: voorbeelden van de verzamelde GIS data binnen de projectzone



Figuur 3-4 Voorbeeld projectzone en GIS data bij 'niet lineair' tracé



(a) GIS shapefiles CBS (inhabitants) and BRT (road info)

(b) GIS shapefiles AVG Groundwater level and BAG (buildings)

### 3.4 Opschonen Database

In totaal zijn er 79 variabelen ("Longlist" Bijlage A) waarvan aan de voorkant verwacht werd, op basis van literatuur en praktijk ervaring van de partners, dat deze invloed zouden kunnen hebben op de aanlegkosten. Dit is de ruwe database.

Na alle data te hebben verzameld is de database onderzocht en opgeschoond op basis van de volgende zaken:

- Variabelen die niet bekend zijn voorafgaand aan de uitvoering (zoals werkelijke projectduur). *Deze variabelen hebben geen voorspellend karakter hebben. Met andere woorden, deze zijn enkel "post – aanleg" bekend. Voorbeeld is de werkelijke duur. Deze variabele kan niet gebruikt worden voor het voorspellen van de aanlegkosten omdat dit vooraf niet bekend is, geplande duur wel.*

- Variabelen met maar 1 of 2 datapunten (te weinig variatie). Hierdoor is bijvoorbeeld de variabele “meekoppelen” afgevallen. Er waren te weinig projecten waar dit speelde.
- Foutieve gegevens (bijvoorbeeld negatieve waarden waar dit niet kan).
- Een of meerdere variabelen zijn gebruikt voor het creëren van een nieuwe variabele (bijvoorbeeld startjaar en startmaand vormen gezamenlijk de nieuwe variabele “tjdstempel”)
- Een of meerdere variabelen zijn opgevraagd ter controle, ten behoeve van “het appels met appels vergelijken”. Bijvoorbeeld, de kosten van afleversets en het aantal warmtestations, om er zeker van te zijn dat deze onderdelen niet meegenomen worden in de analyse omdat dit buiten de onderzoeksscope (1.3) valt.

Deze opschoning heeft geleid tot de uiteindelijk te gebruiken database en betreft daarmee de “shortlist” aan variabelen zoals eerder gepresenteerd in 3.1. Deze “shortlist” database is het startpunt van de kwantitatieve analyse en wordt in de volgende paragraaf en Bijlage A omschreven. In bijlage A is ook te zien welke variabelen wel waren uitgevraagd/geinventariseerd maar welke zijn afgevallen van de shortlist vanwege een van bovenstaande redenen.

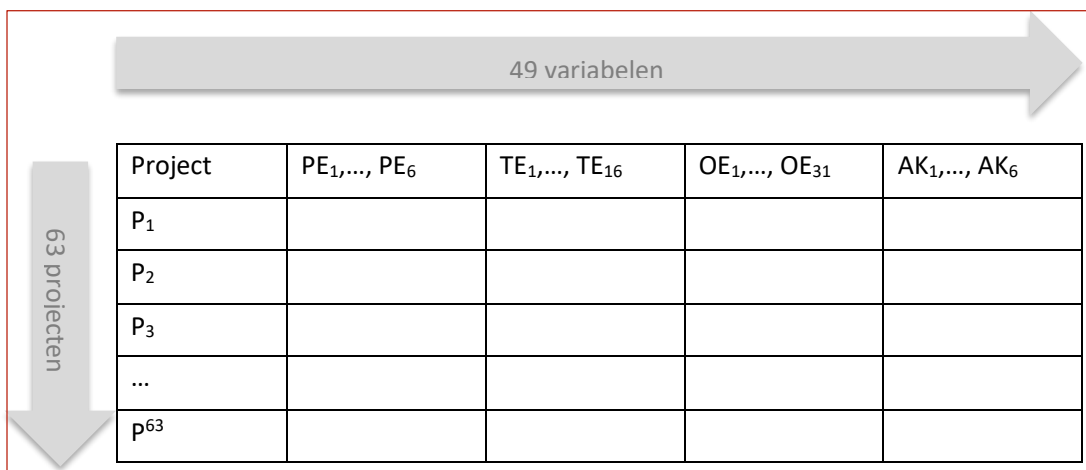
### 3.5 Omschrijving Database

De opgestelde database is enkel beschikbaar voor de projectmedewerkers van Deltares en valt onder de afgesproken NDA (non-disclosure agreement) tussen Deltares en de warmtebedrijven die hun data hebben aangeleverd. De database is, vanwege de bedrijfsgevoelige informatie, niet openbaar beschikbaar.

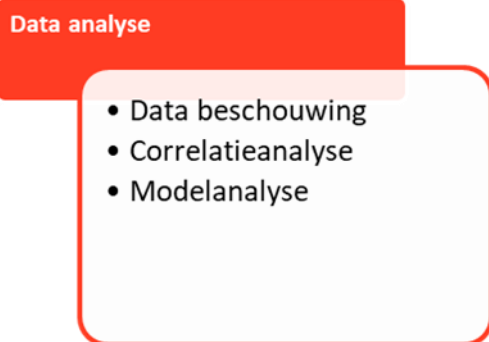
De door de warmtebedrijven aangeleverde data, opgehaald door middel van de opgestelde template, is gekoppeld aan de data die verzameld is uit nationale GIS databases en samengevoegd in de shortlist database. De opgeschoonde dataset bestaat uit 63 projecten en 49 variabelen. Alle variabelen behoren tot een van de vier eerder gepresenteerde categorieën; project eigenschappen (PE), technische eigenschappen (TE), omgevingseigenschappen (OE), en aanlegkosten (AK).

Figuur 3-5 geeft weer hoe de database eruit ziet. Voor een volledige omschrijving van de variabelen en de bron waarvan deze afkomstig zijn zie bijlage A.

Figuur 3-5 Database structuur en omvang met 63 projecten, 5 projecteigenschappen (PE), 13 technische eigenschappen (TE), 31 omgevingseigenschappen (OE), en 6 onderverdelingen in aanlegkosten (AK).



# 4 Data Analyse



## 4.1 Data beschouwing

De eerste stap in de data analyse is het beschouwen van de beschikbare data set. Denk bijvoorbeeld aan de vragen *hoe ziet de spreiding er uit, hoeveel data punten zijn er van verschillende variabelen en zien we bepaalde patronen?*

Belangrijk om vooraf nogmaals te benadrukken is dat de constatering en uitspraken in de volgende paragrafen enkel gaan over de dataset die voor deze analyse beschikbaar is en niet geïnterpreteerd kan worden als algemene waarheden voor alle warmtenet aanleg projecten en alle bedrijven.

### 4.1.1 Aanlegkosten (totaal en per meter)

Figuur 4-1 tot en met Figuur 4-7 geven statistische en visuele weergaven van de beschikbare data over de aanlegkosten (de response variabelen). Zie voor toelichting nogmaals Figuur 1-3 en Figuur 3.1. Daarmee krijgen we inzicht in de financiële omvang van de aanlegkosten van de projecten in onze database. Te zien is dat voor een groot aantal projecten de materiaalkosten niet beschikbaar zijn (missing = 18) en dat daarmee ook de aanlegkosten terreinleidingen, waar materiaalkosten een onderdeel van is veel ontbrekende waarden heeft.

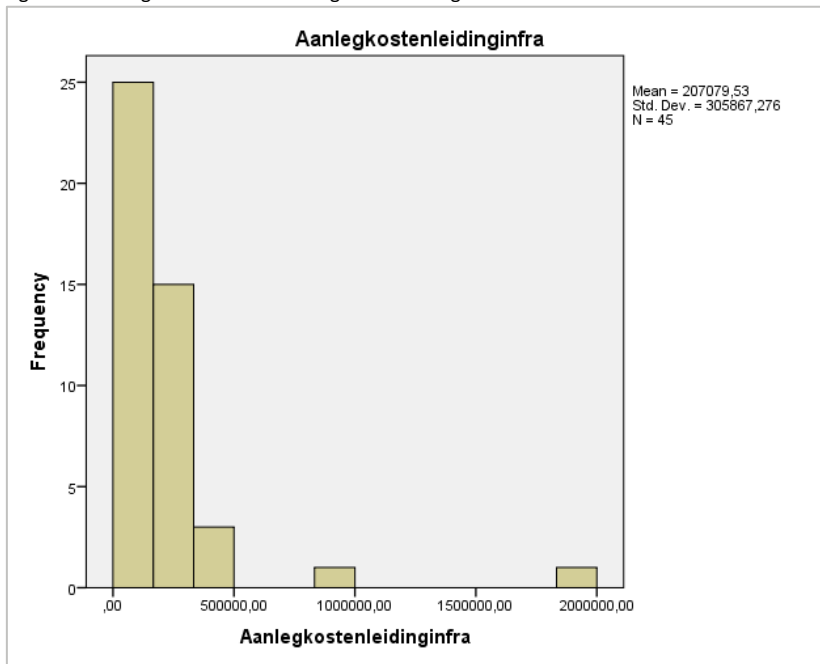
Het merendeel van de projecten hebben aanlegkosten voor de leidinginfra onder de €500,000 met uitschieters van een project van bijna 1 miljoen en een project van bijna 2 miljoen. Wanneer de per meter prijs van de aanlegkosten leidinginfra bekeken wordt zien we een vergelijkbaar patroon met ook twee duurdere projecten (€10k per meter en €17k per meter). Andere projecten vallen onder de €5000 per meter waarbij de grootste groep projecten zich bevindt rond de €1000 per meter.

Wanneer we ook de statistieken van de uitvoeringskosten per meter en de materiaal kosten per meter bekijken is er ook een grote spreiding zichtbaar; de goedkoopste projecten die €48 per meter aan materiaal kosten en €40 per meter aan uitvoeringskosten hebben tegenover de duurste projecten met €6600 per meter aan materiaalkosten en €9700 per meter aan overige kosten. De histogrammen laten zien dat voor materiaalkosten het grootste deel van de projecten onder de €1000 per meter en voor de uitvoeringskosten het merendeel van de projecten onder de €2000 per meter vallen.

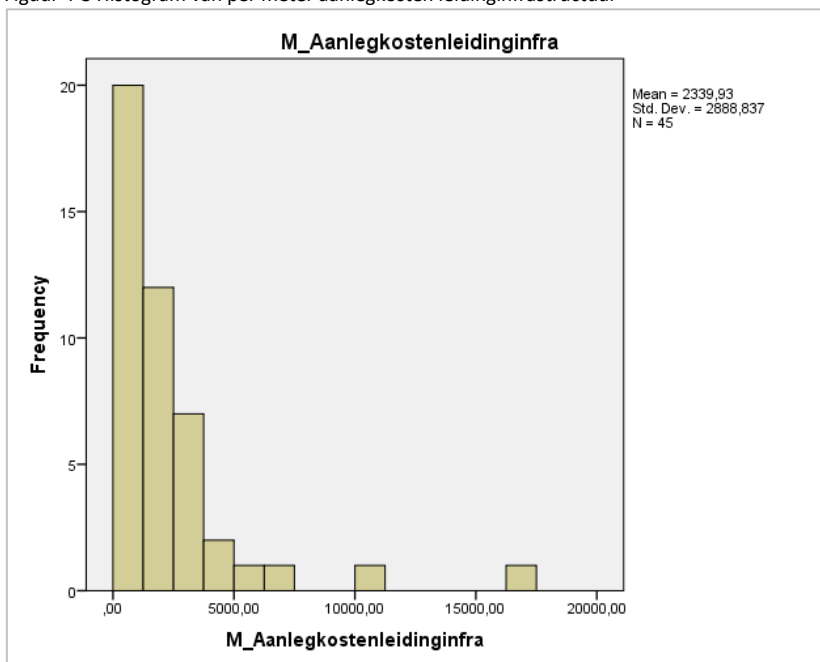
Figuur 4-1 Statistisch overzicht van aanlegkosten; Bovenste rij: naam van variabele. Kolom waarden: aantal (ontbrekende) datapunten, gemiddelde waarde, standaard deviatie van gemiddelde waarde, minimum waarde, maximum waarde

		Statistics					
		Aanlegkosten leidinginfra	Materiaalkosten	Uitvoeringskosten	M_Aanlegkosten leidinginfra	M_Materiaalkosten	M_uitvoeringskosten
N	Valid	45,00	45,00	63,00	45,00	45,00	63,00
	Missing	18,00	18,00	,00	18,00	18,00	,00
Mean		207079,53	60407,97	245327,17	2339,93	774,66	1369,60
Median		140966,00	44120,00	93432,00	1409,53	341,13	821,26
Std. Deviation		305867,28	61776,15	409483,24	2888,84	1357,03	1600,94
Minimum		5685,00	1100,00	3869,00	181,09	48,57	40,56
Maximum		1925500,00	242615,00	1728500,00	16324,93	6595,71	9729,21

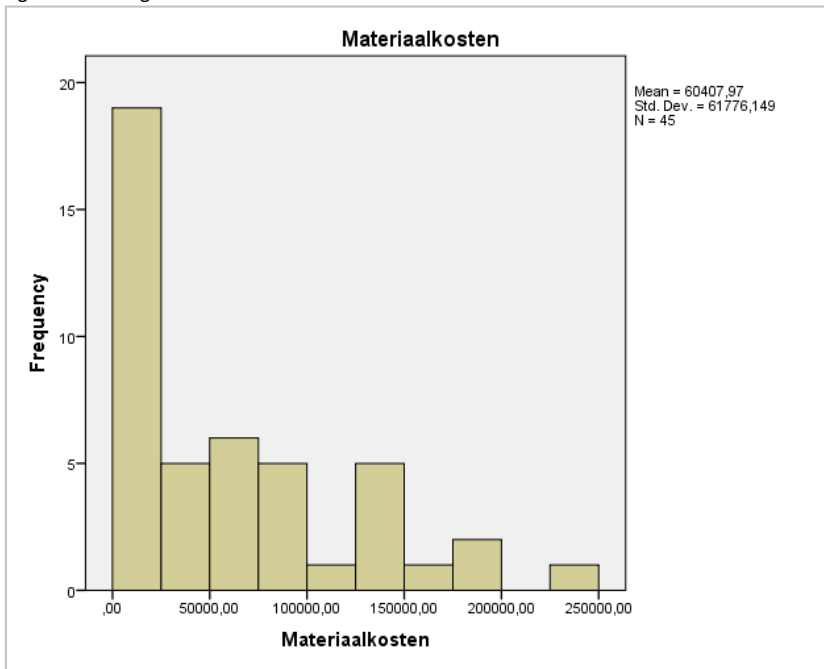
Figuur 4-2 Histogram van totale aanlegkosten leidinginfrastructuur



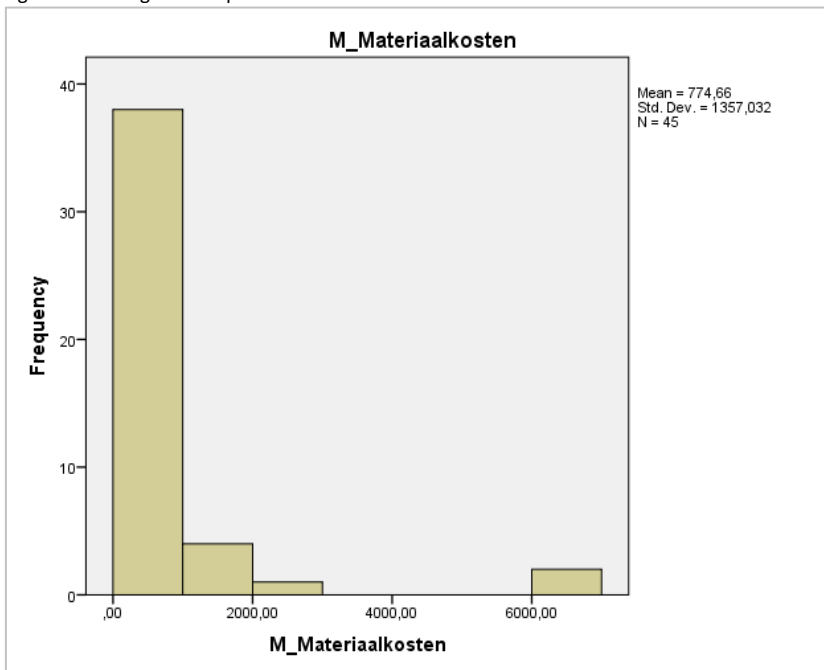
Figuur 4-3 Histogram van per meter aanlegkosten leidinginfrastructuur



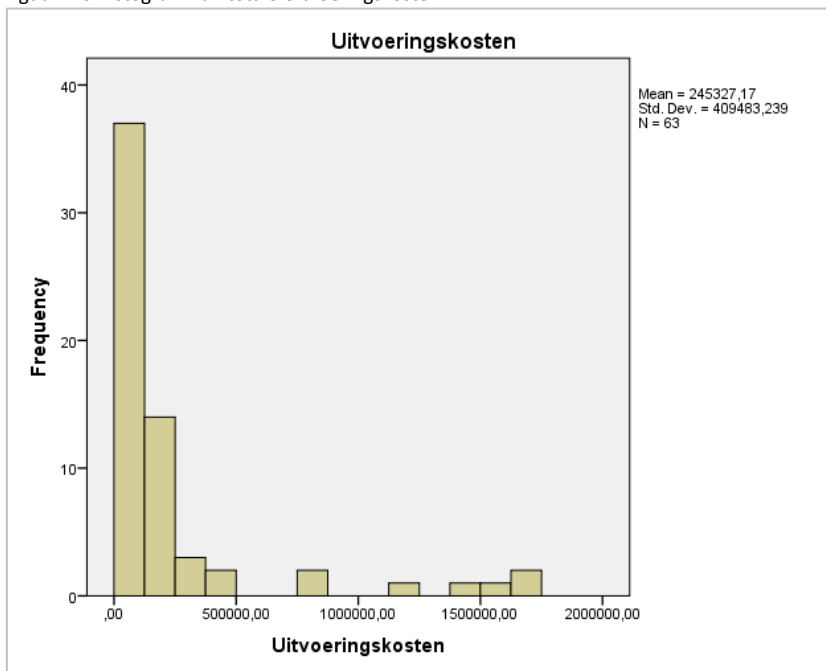
Figuur 4-4 Histogram van totale Materiaalkosten



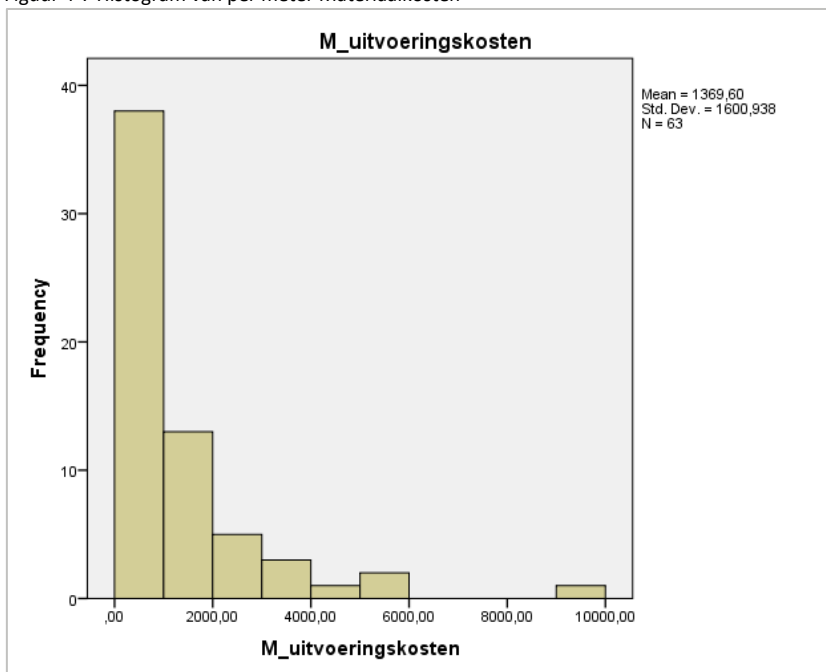
Figuur 4-5 Histogram van per meter Materiaalkosten



Figuur 4-6 Histogram van totale Uitvoeringskosten



Figuur 4-7 Histogram van per meter Materiaalkosten



Uit de hoge standaard deviatie waarden blijkt dat er een flinke spreiding is van de getallen rondom het gemiddelde voor zowel de totale kosten als de kosten per meter. Uit deze gegevens blijkt dan ook dat er een grote bandbreedte van onzekerheid zou zijn wanneer een kosteninschatting gemaakt wordt op basis van enkel de per meter prijs.

Bij de volgende stappen in de statistische analyse wordt alleen gekeken naar de kosten die vallen in de kosten categorie 'uitvoeringskosten' ( Figuur 1-3). Daarbinnen wordt gekeken naar zowel de per meter kosten als de totale kosten. Dit vanwege de ontbrekende data voor de overige kosten categorieën.

## 4.1.2 Projecteigenschappen

Figuur 4-8 tot en met Figuur 4-9 geven een statistische en visuele weergave van de beschikbare data van de projecteigenschappen (een deel van de verklarende variabelen). Daarmee wordt inzicht gegeven in de eigenschappen van de projecten in onze database.

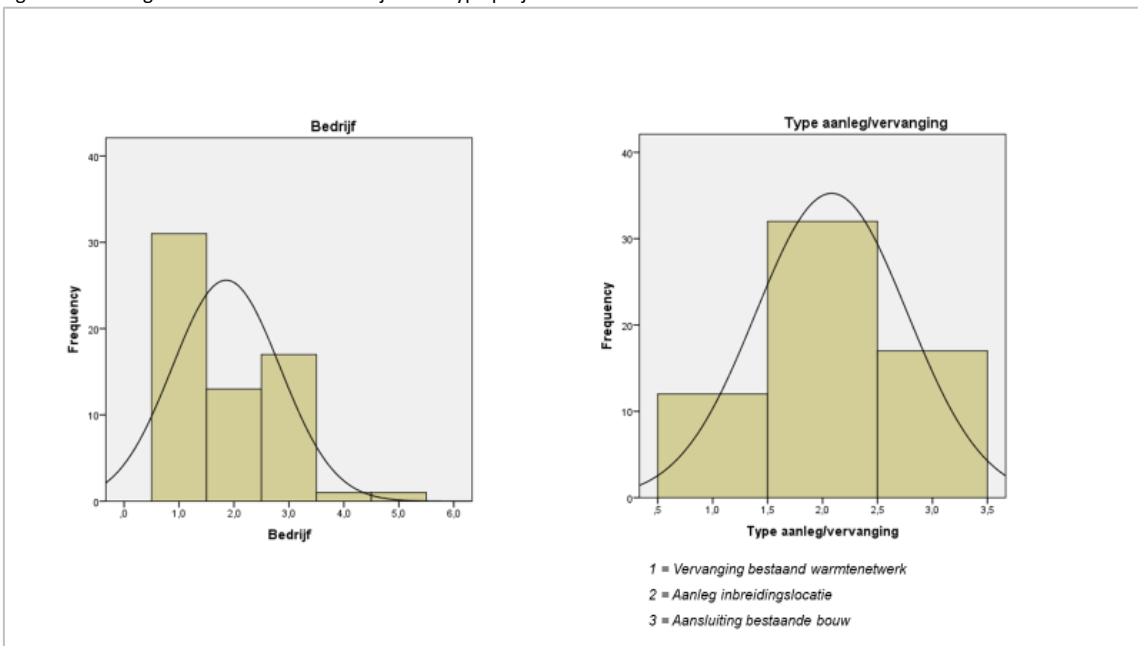
Te zien is dat er data van 5 verschillende bedrijven data is opgenomen in de database. 97% van de projecten zijn afkomstig van 3 bedrijven gezamenlijk. Deze blijven anoniem. Het merendeel van de projecten betreft het type aanleg inbreidingslocatie, en betreft daarnaast bovenal projecten waar enkel de uitvoering gecontracteerd is. Het meest recente project uit de database is gestart in aug 2020 (tijdstempel 44044) en het oudste project is gestart in januari 2015 (tijdstempel 42005).

De geplande duur van de projecten in de database, ook een indicatie van de omvang (op voorhand verwachte omvang, niet de werkelijke), loopt van slechts enkele dagen (min =4) tot een maximum van 611 dagen. Meer dan de helft, 65%, van de projecten in de database hadden een verwachte duur van onder de 200 dagen<sup>2</sup>.

Figuur 4-8 Statistisch overzicht van projecteigenschappen; Bovenste rij: naam van variabele. Kolom waarden: aantal (ontbrekende) datapunten, gemiddelde waarde, standaard deviatie van gemiddelde waarde, minimum waarde, maximum waarde.

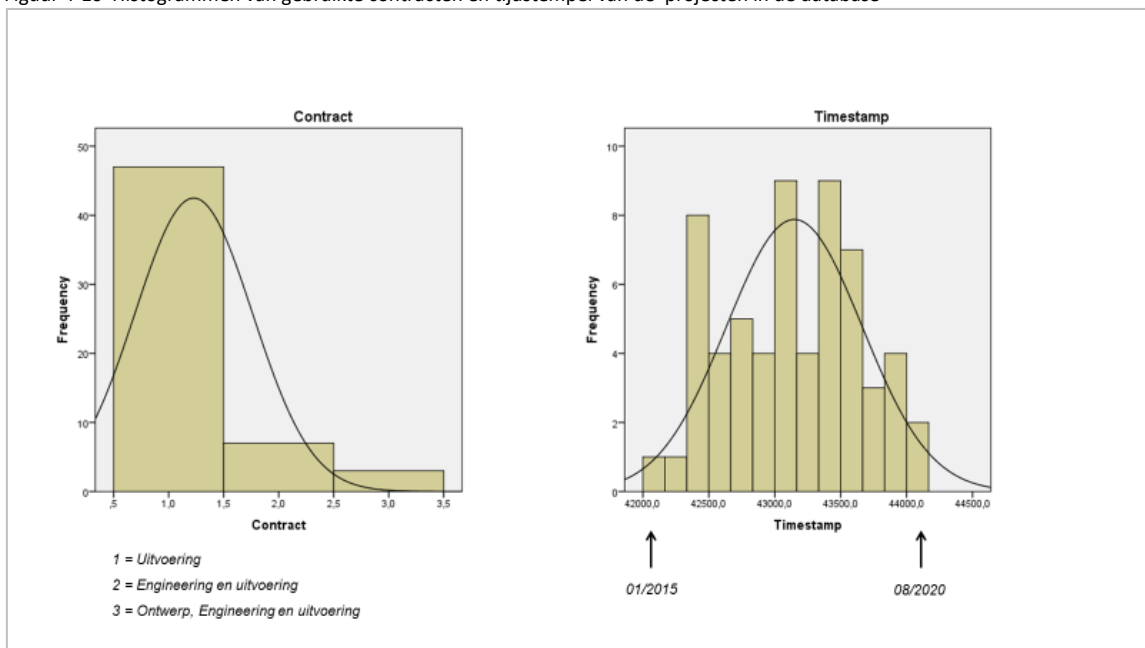
		Statistics				
		Bedrijf	Type	Contract	Geplande duur	Tijdstempel
N	Valid	63,00	61,00	57,00	37,00	61,00
	Missing	,00	2,00	6,00	26,00	2,00
Mean		1,86	2,08	1,23	170,73	43148,21
Median		2,00	2,00	1,00	117,00	43160,00
Std. Deviation		,98	,69	,54	155,66	514,76
Minimum		1,00	1,00	1,00	4,00	42005,00
Maximum		5,00	3,00	3,00	611,00	44044,00

Figuur 4-9 Histogrammen van warmtebedrijven en type projecten in de database

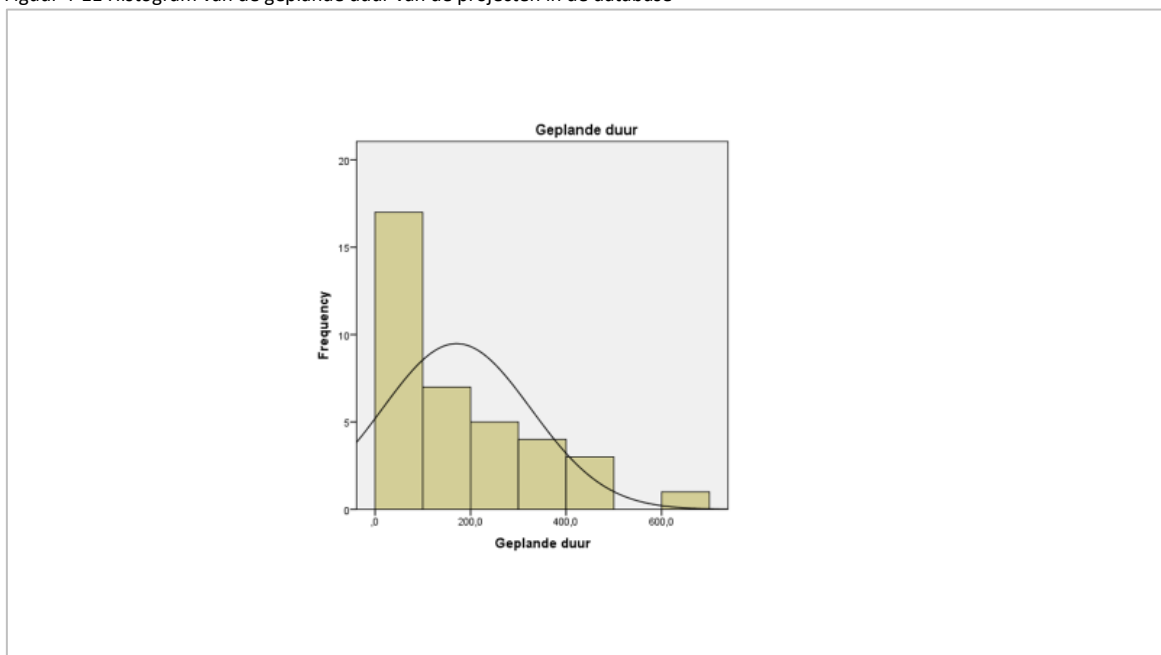


<sup>2</sup> Wanneer we de duur per meter berekenen zien we een max van 22,6 dagen per meter, en een min van 0,08 dagen per meter (iets meer dan 1 uur van een 8uur durende werkdag).

Figuur 4-10 Histogrammen van gebruikte contracten en tijdstempel van de projecten in de database



Figuur 4-11 Histogram van de geplande duur van de projecten in de database



#### 4.1.2.1 Projecteigenschap "Bedrijf"

Er zijn drie belangrijke redenen om de variabele "bedrijf" niet verder mee te nemen in de kwantitatieve vervolganalyse;

- i. Vanwege de commerciële sensitiviteit hiervan. In overeenstemming met de warmtebedrijven (NDA) om vertrouwelijk met de aangeleverde, gevoelige data, om te gaan, mag het niet mogelijk zijn om data te herleiden naar individuele bedrijven of om



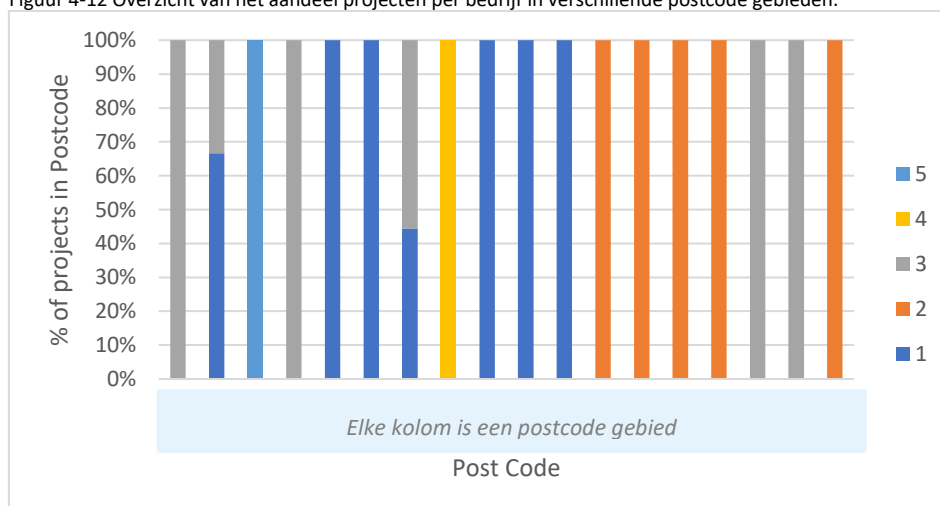
resultaten, bijvoorbeeld kosten kentallen, naar verschillende bedrijven af te leiden. Het aanhouden van deze variabele levert hierdoor geen informatie op die gebruikt mag worden.

- ii. Omdat verschillende project partners hebben aangegeven dat de aanlegkosten niet significant beïnvloed worden door welk warmtebedrijf het project uitvoert. Wanneer exact hetzelfde project uitgevoerd zou worden door verschillende partijen zal dit leiden tot (bijna) gelijke aanlegkosten. Daarmee beargumenteren de projectpartners, dat “bedrijf” geen verklarende factor is voor de verschillend in aanlegkosten.
- iii. Omdat dit geen inzichten oplevert waar naar gehandeld kan worden bij toekomstige warmtenet aanlegprojecten. Het is immers niet op voorhand bekend aan welke partij het project gegund zal worden, noch kunnen er partijen worden uitgesloten.

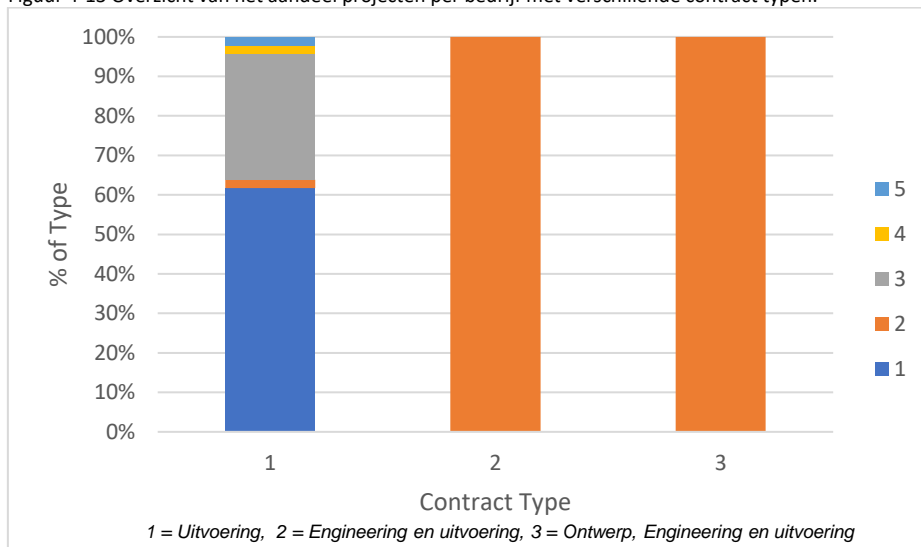
Uit een eerste correlatie analyse en modelopbouw leek het echter dat de aanlegkosten wel bedrijfsafhankelijk waren en dat er dus wel een significante correlatie bestond tussen “bedrijf” en de aanlegkosten. Gegeven bovenstaand argument 2 is het goed mogelijk dat dit kwam door onderliggende karakteristieken van de projecten van de verschillende bedrijven, zoals bijvoorbeeld dat veel of alle projecten van bedrijf X in dezelfde gemeente hadden plaatsgevonden, of dat juist alle projecten van bedrijf X de grotere projecten betrof. Dit gaf aanleiding om de variabele “bedrijf” nader te onderzoeken alvorens deze in de vervolg analyse uit te sluiten.

Figuur 4-12, Figuur 4-13, Figuur 4-14 en Figuur 4-15 laten zien hoe de projectportfolio in onze database van de verschillende bedrijven verdeeld is over postcode, type contract, project grootte, en type project. Wat betreft postcode gebied, lijkt er weinig overlap te zijn van bedrijven die projecten uitvoeren in hetzelfde postcode gebied enkel in twee postcode gebieden hebben twee bedrijven projecten uitgevoerd. Wat betreft de gecontracteerde werkzaamheden is er ook een duidelijk onderscheid in de voorkeur (of mogelijkheden) van bedrijven. Bij vier van de vijf bedrijven is 100% van hun portfolio alleen uitvoering gecontracteerd, en bij het vijfde bedrijf bijna 100% uitvoering + engineering (+ontwerp). Dezelfde trend is te zien wanneer we kijken naar de grootte van de projecten in de project portfolio's. Een enkel bedrijf domineert in de kleinere aanlegprojecten, terwijl bij de grotere projecten treffen we een mengeling van alle bedrijven. Ditzelfde patroon is te zien bij de type projecten die uitgevoerd worden door de verschillende bedrijven.

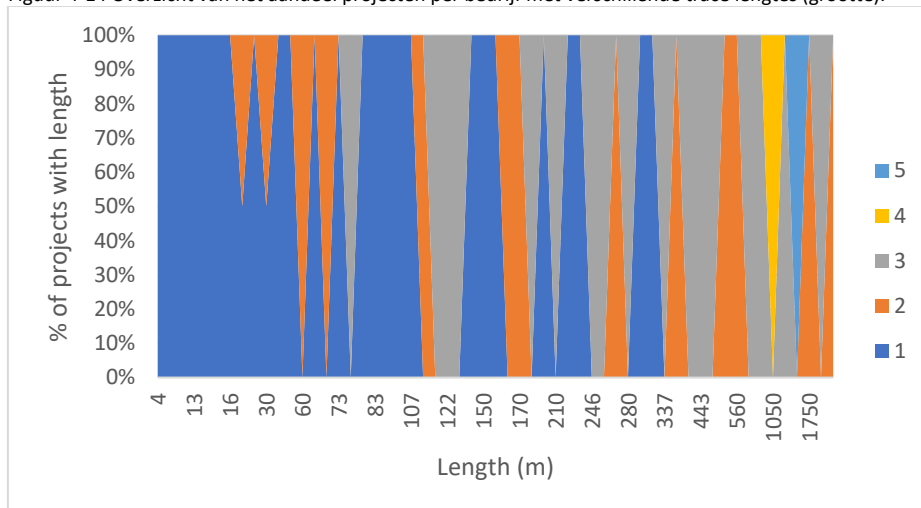
Figuur 4-12 Overzicht van het aandeel projecten per bedrijf in verschillende postcode gebieden.



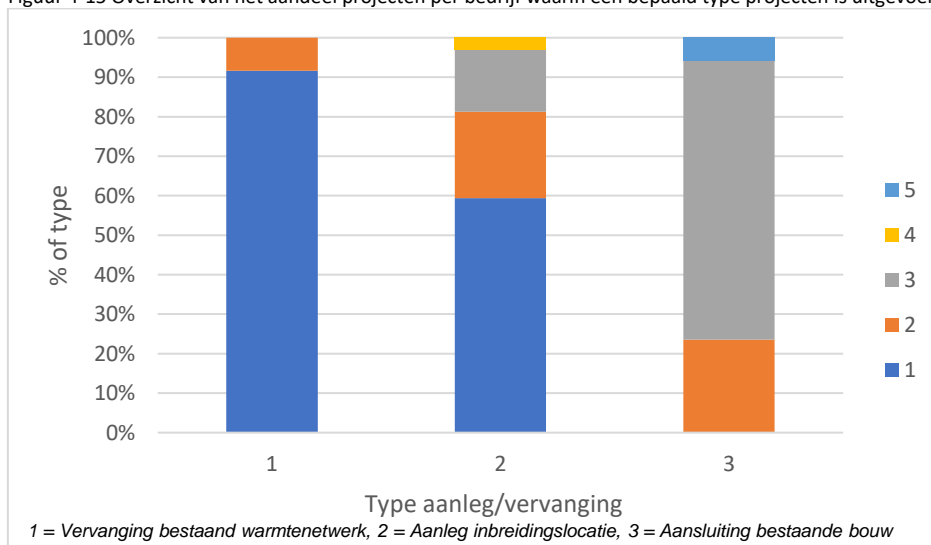
Figuur 4-13 Overzicht van het aandeel projecten per bedrijf met verschillende contract typen.



Figuur 4-14 Overzicht van het aandeel projecten per bedrijf met verschillende tracé lengtes (grootte).



Figuur 4-15 Overzicht van het aandeel projecten per bedrijf waarin een bepaald type projecten is uitgevoerd.



### 4.1.3 Technische eigenschappen

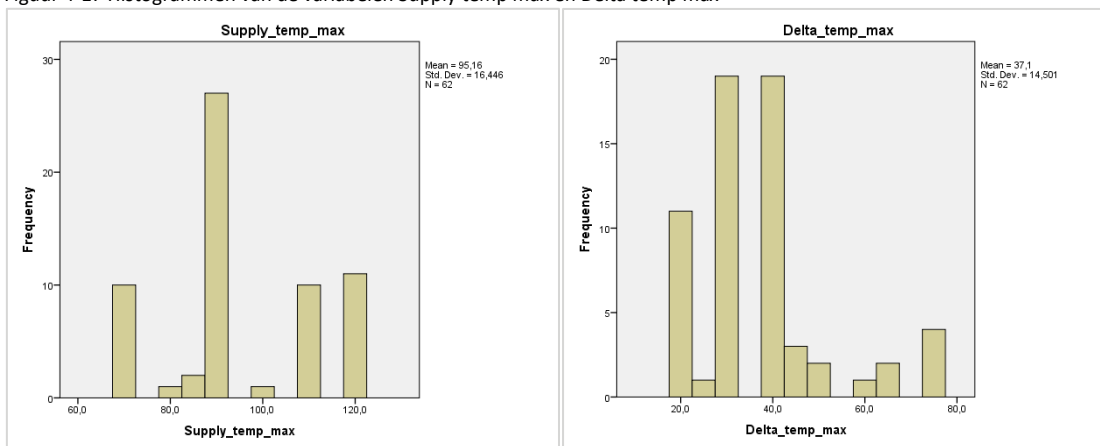
Figuur 4-16 tot en met Figuur 4-23 geven een statistische en visuele weergave van de beschikbare data van de technische eigenschappen (een deel van de verklarende variabelen). Daarmee wordt inzicht gegeven in de eigenschappen van de projecten in onze database.

Te zien is dat er voor een groot aantal variabelen van de technische eigenschappen veel ontbrekende data is. Het betreft vooral ontbrekende data over het aantal afnemers, aantal aansluitingen, aangelegd vermogen en aangesloten vermogen. Het kleinste project in de data set is een project van 4 meter leiding lengte, en het langste project beslaat 2264meter. Qua temperatuur van het warmtenet betreft het enkel projecten van boven de 70 graden aanvoer temperatuur en daarmee geen midden en laag temperatuur warmtenetten (Textbox 1.1). ruimte op het net is bepaald op basis van het vermogen dat is aangesloten en het vermogen waarop het net is gedimensioneerd. Omdat er bij beide variabelen veel data ontbreekt zijn er voor de variabel ruimte op het net ook veel missende data punten.

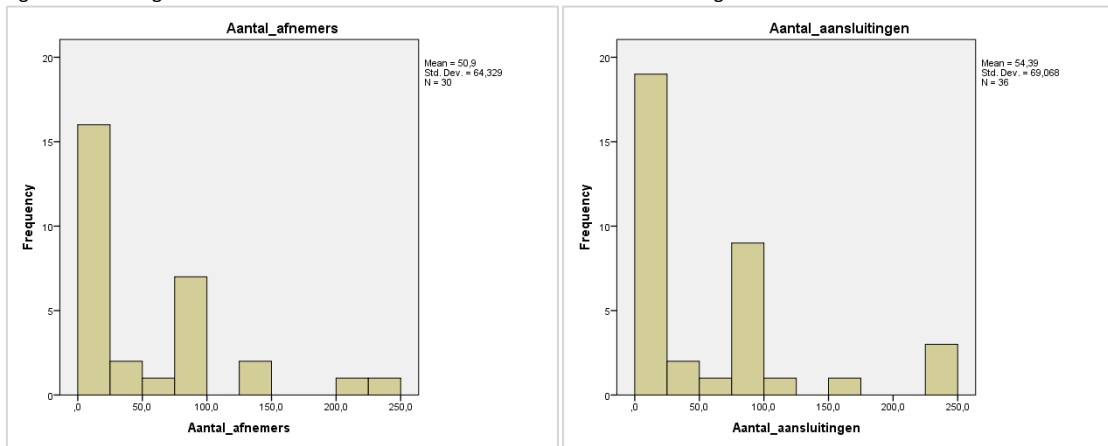
Figuur 4-16 Statistisch overzicht van technische eigenschappen; Bovenste rij: aantal (ontbrekende) datapunten, gemiddelde waarde, standaard deviatie van gemiddelde waarde, minimum waarde, maximum waarde. Kolom: naam van variabele

	N		Mean	Median	Std. Deviation	Minimum	Maximum
	Valid	Missing					
Supply_temp_max	62,00	1,00	95,16	90,00	16,45	70,00	120,00
Delta_temp_max	62,00	1,00	37,10	35,00	14,50	20,00	75,00
Lengte	63,00	,00	321,60	128,00	507,30	4,00	2264,00
Aantal_afnemers	30,00	33,00	50,90	19,00	64,33	1,00	231,00
Aantal_aansluitingen	36,00	27,00	54,39	19,00	69,07	1,00	231,00
Aangesloten_vermogen	33,00	30,00	688,88	450,00	978,92	28,00	5500,00
Aangelegd_vermogen	37,00	26,00	741,86	560,00	946,19	28,00	5500,00
Ruimteophetnet	32,00	31,00	,25	,00	,44	,00	1,00
Dominant_leidingtype	63,00	,00	1,57	1,00	,76	1,00	3,00
Dominant_leidingmateriaal	63,00	,00	1,43	1,00	1,00	1,00	4,00
Leiding_diameter	61,00	2,00	103,48	70,00	80,84	25,00	315,00
Methode_opensleuf	63,00	,00	,79	1,00	,41	,00	1,00
Diepte_opensleuf	50,00	13,00	1,17	1,00	,37	,80	2,00

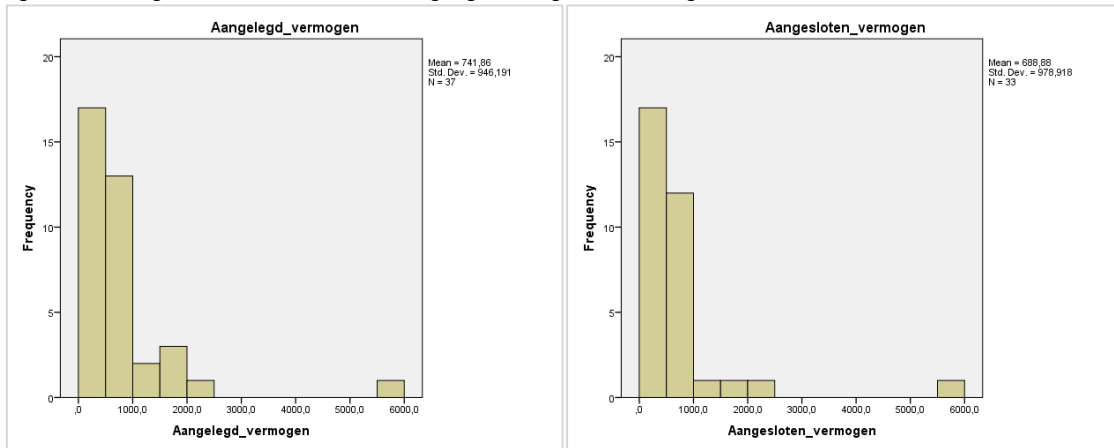
Figuur 4-17 Histogrammen van de variabelen Supply temp max en Delta temp max



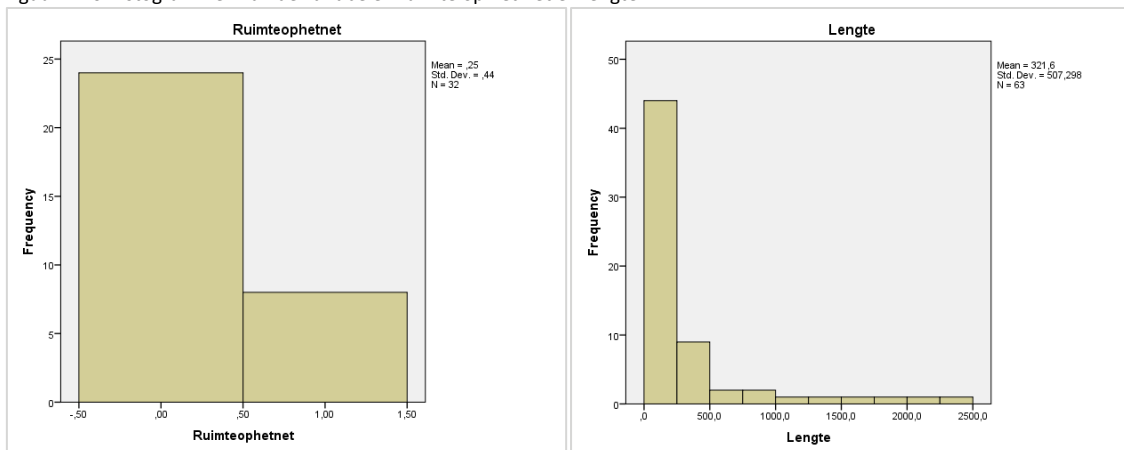
Figuur 4-18 Histogrammen van de variabelen aantal afnemers en aantal aansluitingen



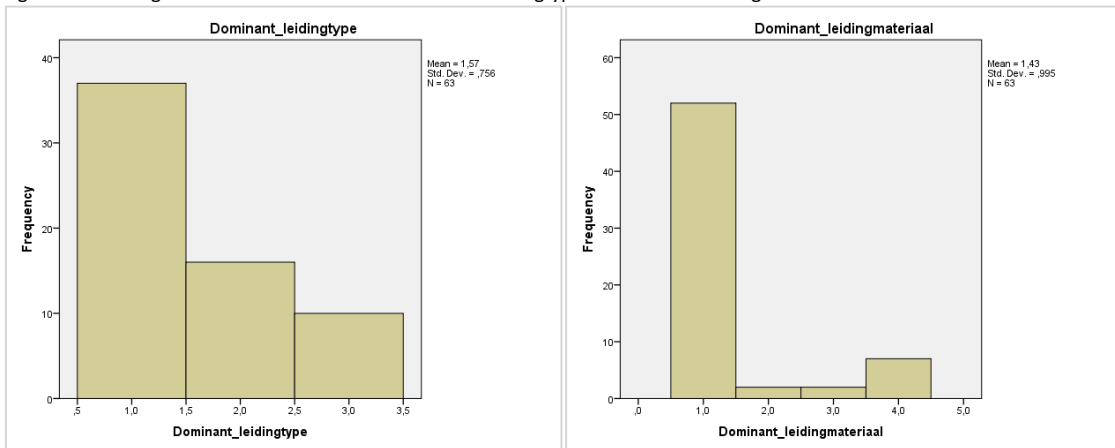
Figuur 4-19 Histogrammen van de variabelen aangelegd en aangesloten vermogen



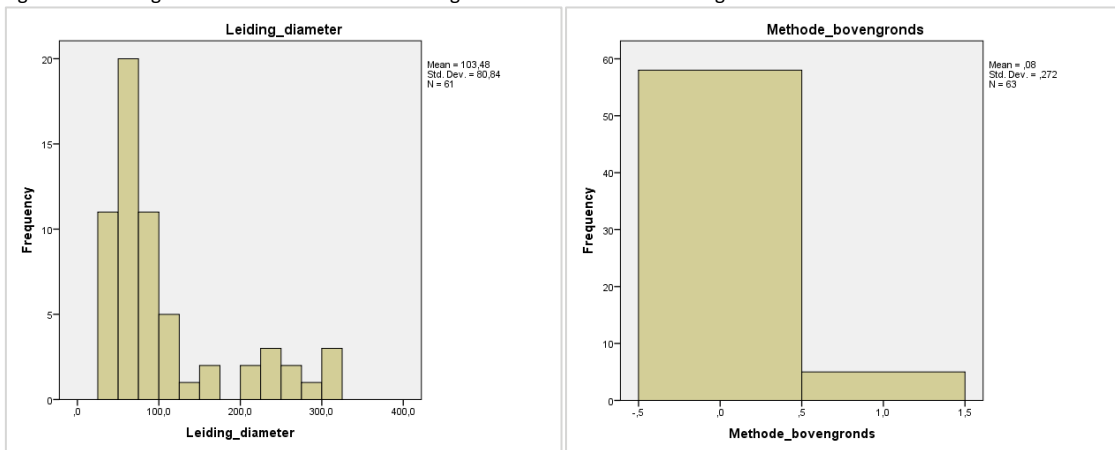
Figuur 4-20 Histogrammen van de variabelen ruimte op het net en lengte



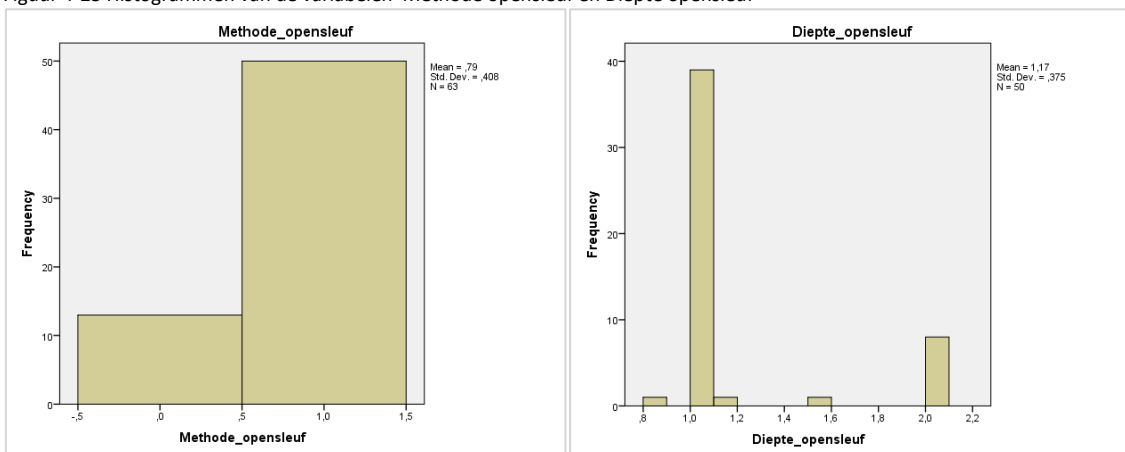
Figuur 4-21 Histogrammen van de variabelen dominant leidingtype en dominant leiding materiaal



Figuur 4-22 Histogrammen van de variabelen leiding diameter en Methode bovengronds



Figuur 4-23 Histogrammen van de variabelen Methode opensleuf en Diepte opensleuf



#### 4.1.4 Omgevingseigenschappen

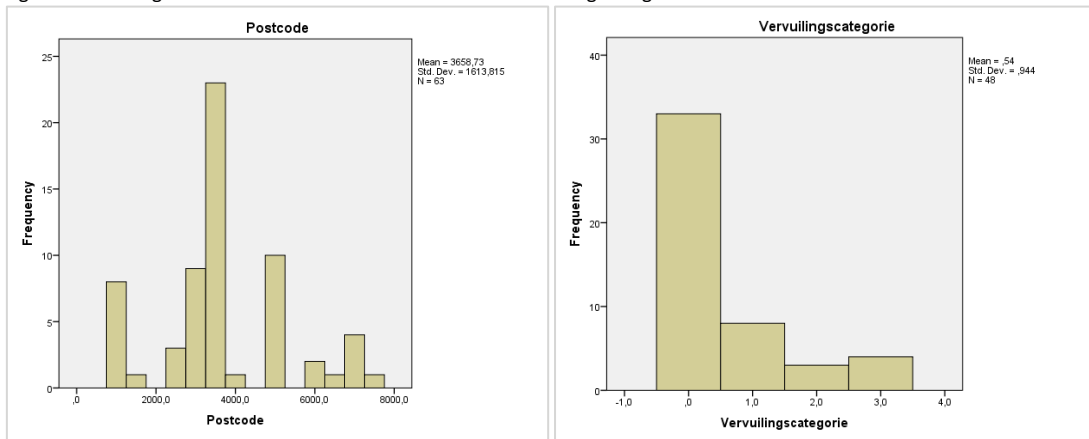
Figuur 4-24 tot en met Figuur 4-39 geven een statistische en visuele weergave van de beschikbare data van de omgevingseigenschappen (een deel van de verklarende variabelen). Daarmee wordt inzicht gegeven in de eigenschappen van de projecten in onze database.

Het merendeel van deze variabelen is afkomstig uit GIS bestanden. Er is minder sprake van missende data dan bij de vorige variabelen. De bijbehorende histogrammen zijn hieronder ook toegevoegd.

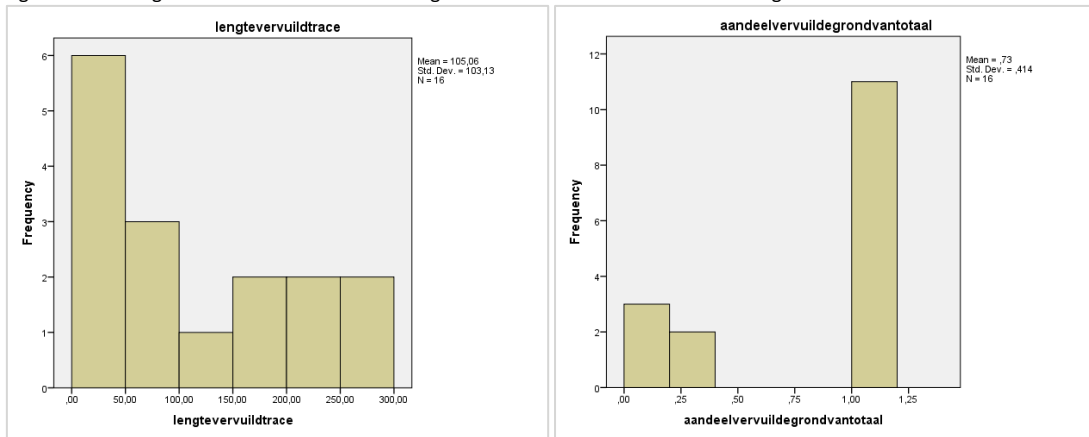
Figuur 4-24 Statisch overzicht van omgevingseigenschappen; Bovenste rij: aantal (ontbrekende) datapunten, gemiddelde waarde, standaard deviatie van gemiddelde waarde, minimum waarde, maximum waarde . Kolom: naam van variabel

	N		Mean	Median	Std. Deviation	Minimum	Maximum
	Valid	Missing					
Postcode	63,00	,00	3658,73	3500,00	1613,81	1000,00	7500,00
Vervuilingscategorie	48,00	15,00	,54	,00	,94	,00	3,00
lengtevervuiltrace	16,00	47,00	105,06	79,00	103,13	3,00	295,00
aandeelvervuildegroondv antotaal	16,00	47,00	,73	1,00	,41	,02	1,00
Bemaling	56,00	7,00	,50	,50	,50	,00	1,00
Grondverbetering	56,00	7,00	,70	1,00	,46	,00	1,00
Wegtype_Hoofdweg	63,00	,00	,05	,00	,21	,00	1,00
Wegtype_regionaleweg	63,00	,00	,21	,00	,41	,00	1,00
Wegtype_lokaleweg	63,00	,00	,37	,00	,49	,00	1,00
Wegbreedte_boven7m	63,00	,00	,43	,00	,50	,00	1,00
Wegbreedte_4tot7m	63,00	,00	,16	,00	,37	,00	1,00
Wegbreedte_2tot4m	62,00	1,00	1,00	1,00	,00	1,00	1,00
Wegbreedte_tot2m	63,00	,00	,19	,00	,40	,00	1,00
Wegbreedte	63,00	,00	9,17	8,00	4,73	1,00	35,00
Voetafdruk_Weg	63,00	,00	,26	,24	,11	,09	,59
Voetafdruk_Wegtype_ver binding	62,00	1,00	,80	,87	,19	,19	1,00
Voetafdruk_Wegtype_krui sing	57,00	6,00	,10	,09	,06	,01	,30
Voetafdruk_Wegtype_ove rigverkeersgebied	25,00	38,00	,31	,30	,28	,00	1,00
Voetafdruk_gebouw	61,00	2,00	,22	,20	,13	,04	,54
Bouwjaar	61,00	2,00	1961,13	1977,00	58,64	1734,00	2018,00
Standaarddeviatie_Bouwj aar	61,00	2,00	32,95	18,00	48,22	,00	206,00
Afstandtussengebouwen weg	61,00	2,00	10,36	9,00	5,63	3,00	31,00
Inwonersdichtheid	63,00	,00	77,63	73,00	53,75	,00	213,00
Grondwater_niveau	63,00	,00	1,79	1,89	,48	,29	2,25
Woongebied	63,00	,00	,46	,00	,50	,00	1,00
Spoorlijn	63,00	,00	,08	,00	,27	,00	1,00
Waterlichaam	63,00	,00	,06	,00	,25	,00	1,00
Peilgebied_niveau	44,00	19,00	-6,65	,27	2,28	-6,70	7,35
Voetafdruk_Bomen	63,00	,00	19,68	19,00	11,29	,00	51,00
Trefkans_archeologie	47,00	16,00	1,83	1,75	,84	1,00	3,00

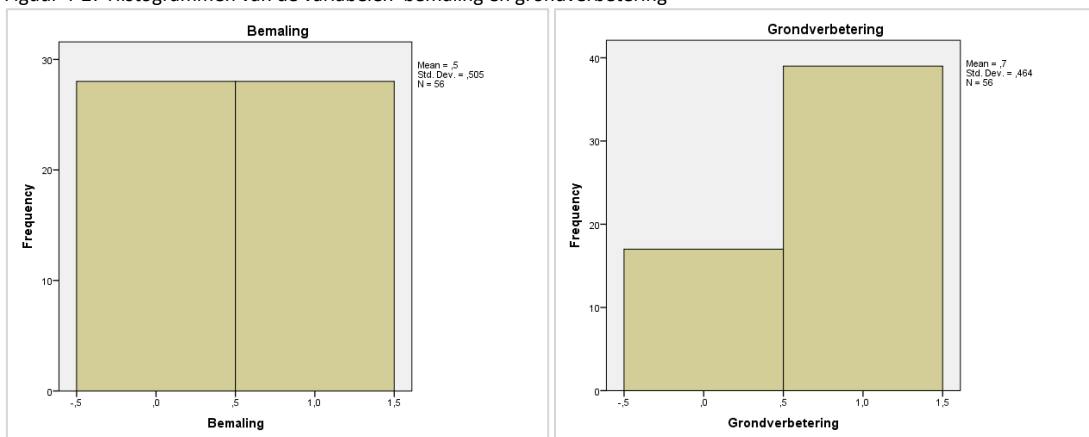
Figuur 4-25 Histogrammen van de variabelen Postcode en Vervuilingscategorie



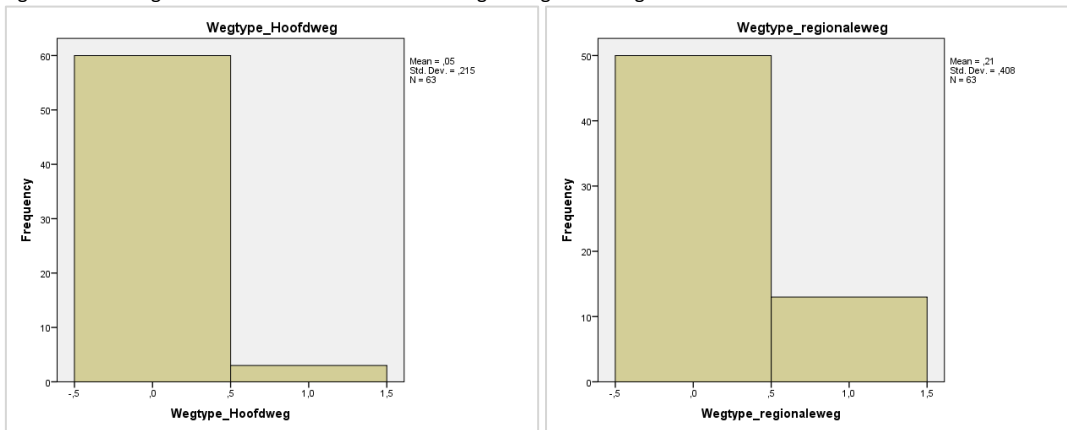
Figuur 4-26 Histogrammen van de variabelen lengte vervuilde trace en aandeel vervuilde grond van totaal



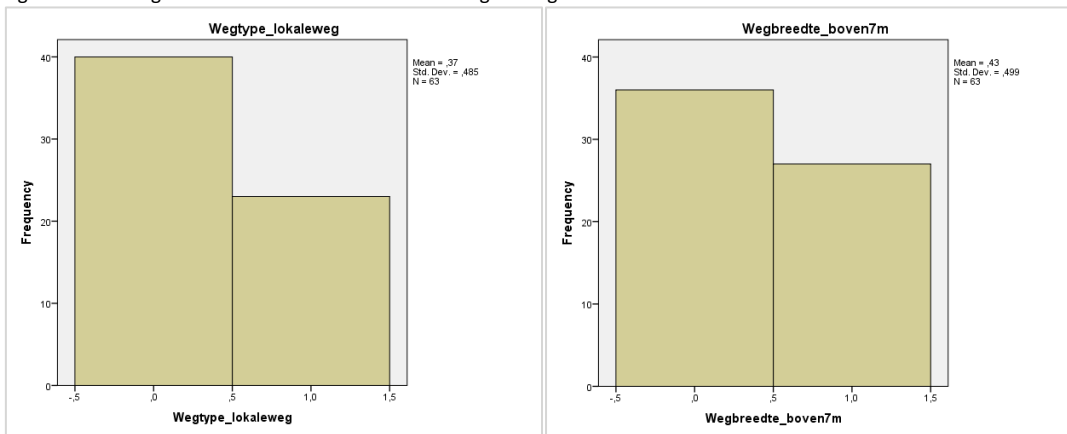
Figuur 4-27 Histogrammen van de variabelen bemaling en grondverbetering



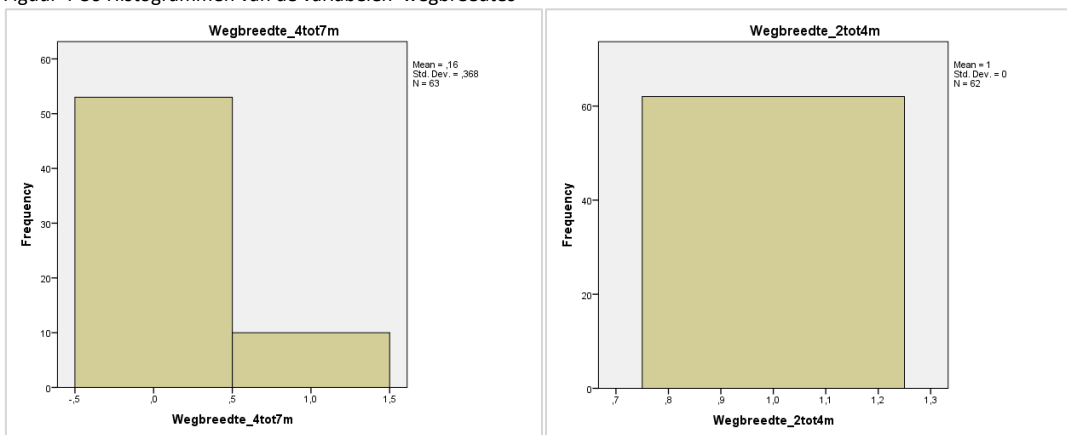
Figuur 4-28 Histogrammen van de variabelen Hoofdweg en Regionale weg



Figuur 4-29 Histogrammen van de variabelen lokale weg en wegbreedte >7m

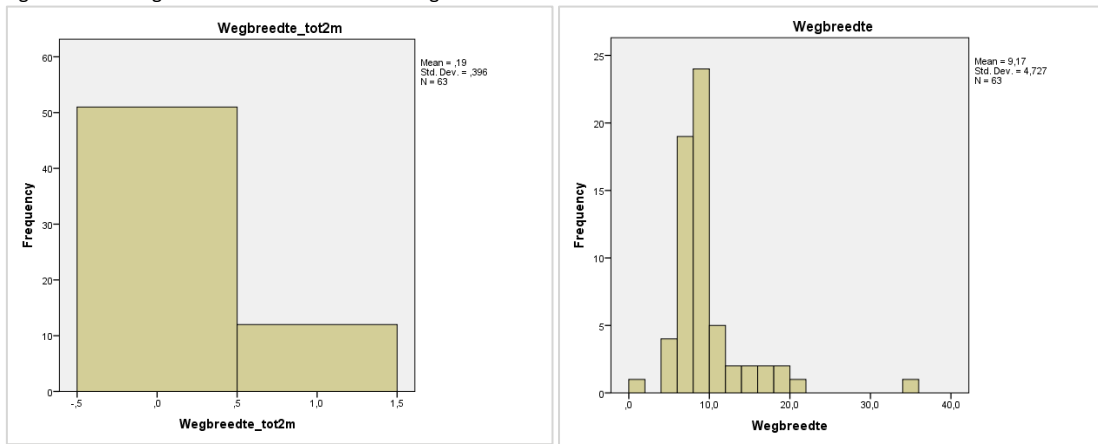


Figuur 4-30 Histogrammen van de variabelen wegbreedtes

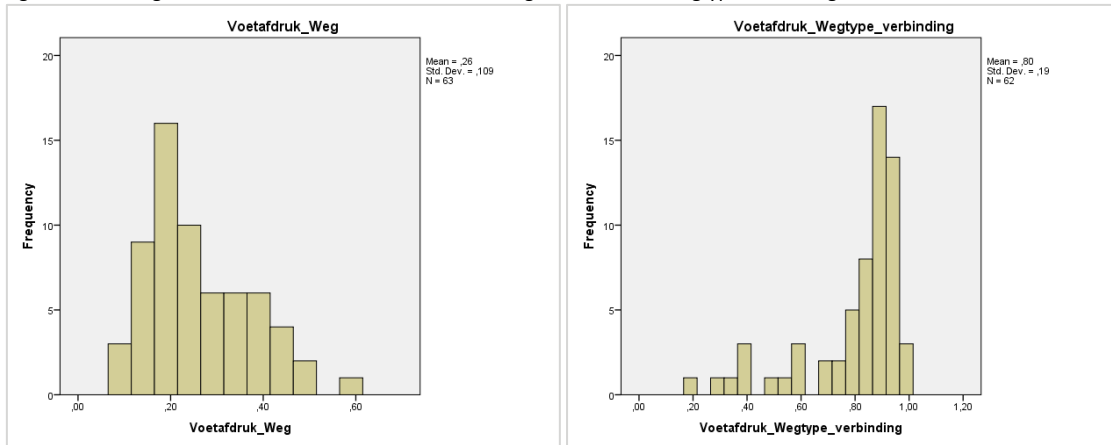




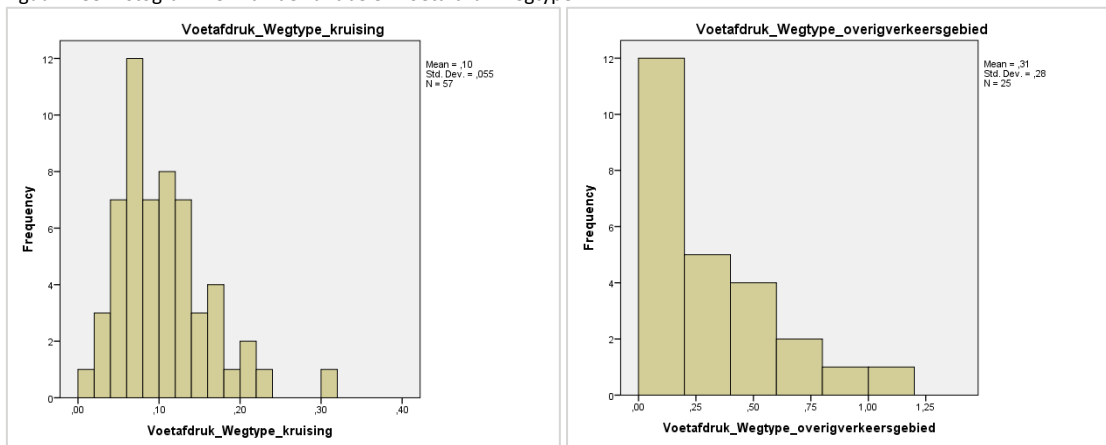
Figuur 4-31 Histogrammen van de variabelen wegbreedtes



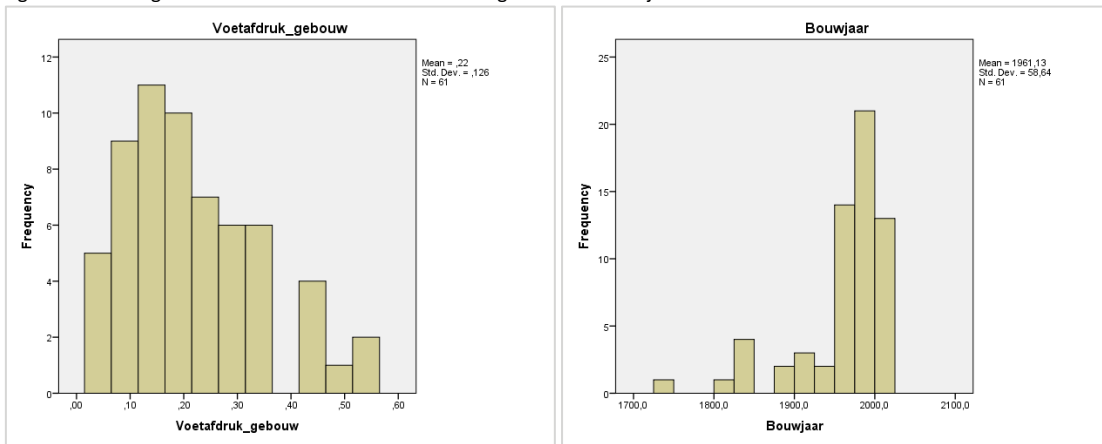
Figuur 4-32 Histogrammen van de variabelen voetafdruk weg en voetafdruk wegtype verbinding



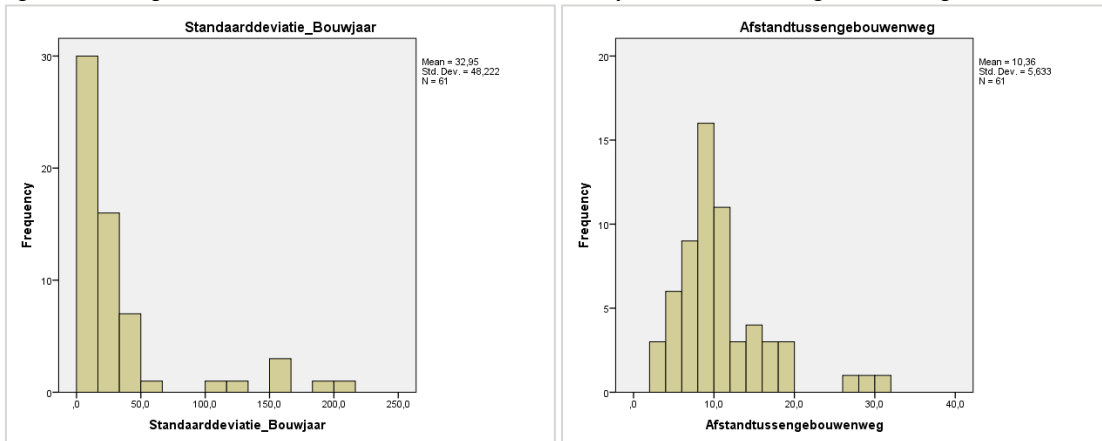
Figuur 4-33 Histogrammen van de variabelen voetafdruk wegtypen



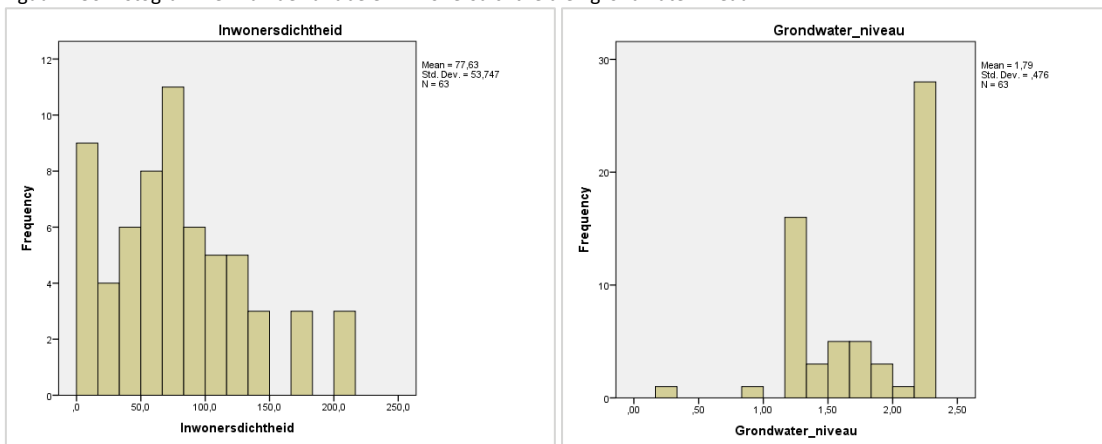
Figuur 4-34 Histogrammen van de variabelen voetafdruk gebouw en bouwjaar



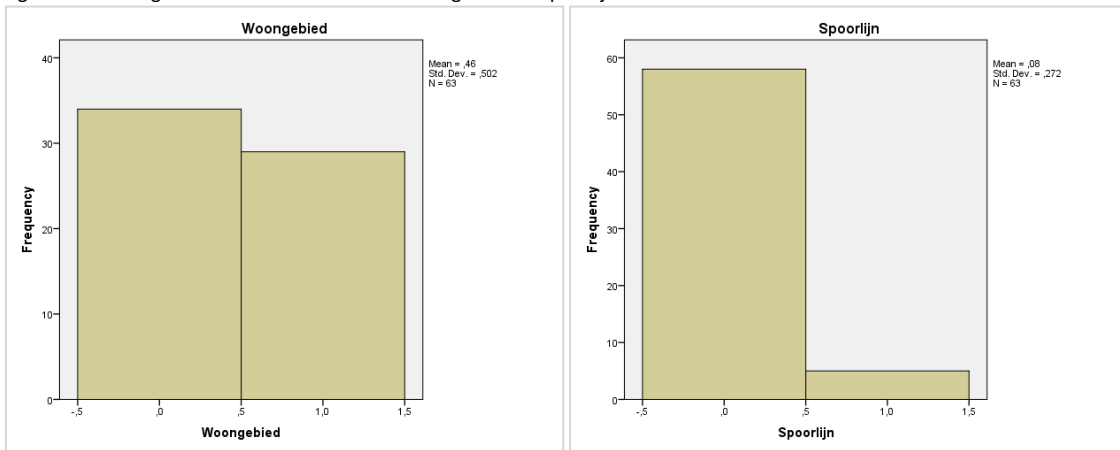
Figuur 4-35 Histogrammen van de variabelen standaard deviatie bouwjaar en afstand tussen gebouw en weg



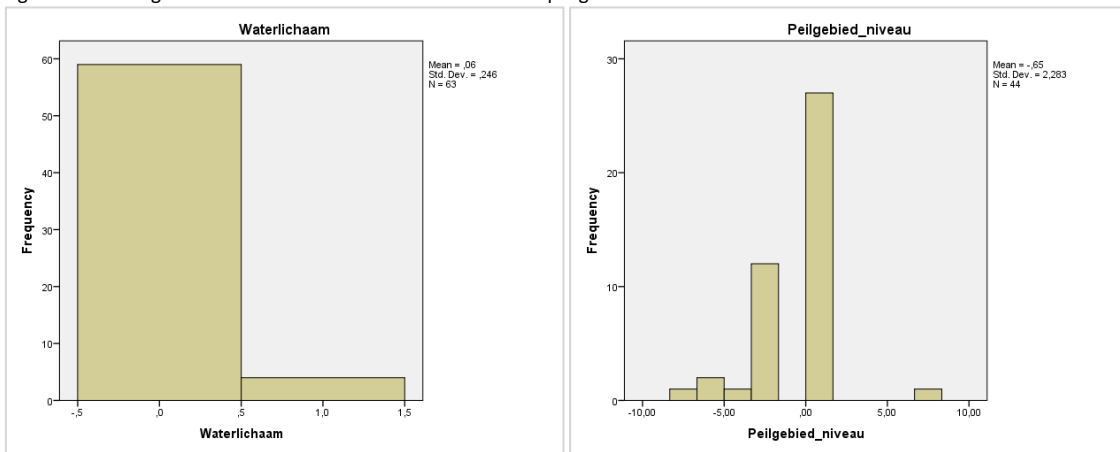
Figuur 4-36 Histogrammen van de variabelen inwonersdichtheid en grondwaterniveau



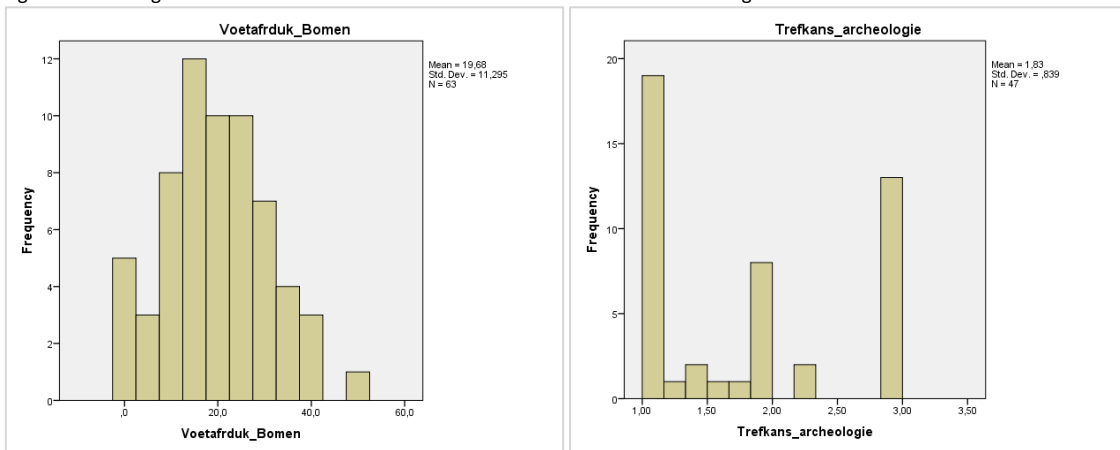
Figuur 4-37 Histogrammen van de variabelen woongebied en spoorlijn



Figuur 4-38 Histogrammen van de variabelen waterlichaam en peilgebied niveau



Figuur 4-39 Histogrammen van de variabelen voetafdruk bomen en trefkans archeologie



## 4.2 Correlatie analyse

Voorafgaand moet benadrukt worden dat deze data set een kleine dataset betreft met veel ontbrekende data punten. Het is dus moeilijk om op basis hiervan sterke, statistisch significante, conclusies te kunnen trekken. Desalniettemin levert het wel inzichten op, welke aangevuld en geïnterpreteerd worden met praktisch expertise van de projectpartners.

De correlatie analyse (ook feature analysis genoemd) is een simpele statistische analyse waarin het verband tussen de variabelen wordt onderzocht. In dit geval het verband tussen aan de ene kant de aanlegkosten van de projecten en aan de andere kant de verschillende eigenschappen van de projecten. Deze analyse wordt gebruikt om de sterkte van een verband tussen variabelen te onderzoeken. Wanneer er sprake is van een correlatie is er een statistisch verband tussen variabelen maar dit betekenen niet dat er een causaal (oorzakelijk) verband is.

De correlatie analyse geeft al een eerste indicatie hoe de variabelen variëren met de aanlegkosten. Volgens gebruikelijke werkwijze wordt er alleen gekeken naar absolute correlatie waarden van boven de 0.2. De correlatie analyse wordt uitgevoerd voor de variabelen Uitvoeringskosten en M\_Uitvoeringskosten. De resultaten worden in Figuur 4-40 weergegeven.

Figuur 4-40 Resultaten correlatie analyse

correlatie met Uitvoeringskosten		correlatie met M_Uitvoeringskosten	
variabel	correlatie-coëfficiënt	variabel	correlatie-coëfficiënt
Aangelegd_vermogen	,679**	Waterlichaam	,473**
Aangesloten_vermogen	,670**	Voetafdruk_gebouw	,340**
Lengte	,649**	Trefkans_archeologie	,295*
Aantal_aansluitingen	,423*	Delta_temp_max	,277*
Aantal_afnemers	,417*	Standaarddeviatie_Bouwjaar	,269*
Leiding_diameter	,383**	Grondwater_niveau	,253*
Type	,278*	Voetafdruk_Wegtype_overigverkeersgebied	,243
Dominant_leidingtype	,275*	Grondverbetering	,233
Bemaling	,274*	Supply_temp_max	,219
Woongebied	,273*	Type	-,203
aandeelvervuildegrondvantaal	,268	Wegbreedte_tot2m	-,207
Dominant_leidingmateriaal	-,210	Bemaling	-,240
Postcode	-,217	Lengte	-,244
Trefkans_archeologie	-,314*	Geplande duur	-,251
Voetafdruk_Wegtype_overigverkeersgebied	-,332	Afstandtussengebouwenweg	-,278*
		Bouwjaar	-,348**
		aandeelvervuildegrondvantaal	-,385

*gemarkeerde velden zijn de variabelen die zowel in de totaal als de per meter kosten een correlatie tonen.*

\* Correlation is significant at the 0.05 level (2-tailed)  
 \*\* Correlation is significant at the 0.01 level (2-tailed)

Positieve waarde geven een positieve relatie aan tussen de betreffende variabele en de kosten variabele. Bijvoorbeeld, in de linker tabel geeft de correlatie coëfficiënt van *Aangelegd\_vermogen\_kW* (,679) aan dat een toename in aangelegd vermogen gepaard gaat met een toename in Uitvoeringskosten. Negatieve waarden geven een negatieve relatie aan tussen de betreffende variabele en de kosten variabele. Bijvoorbeeld, in de rechter tabel, geeft de correlatie coëfficiënt van *Lengte* (-,244) aan dat een toenemende lengte gepaard gaat met een afname in *M\_Uitvoeringskosten*. Hoe hoger de coëfficiënt, hoe sterker de correlatie.

De linker tabel laat zien dat bovenal de eigenschappen die een indicatie geven van de fysieke en technische omvang (lengte, verschillende vermogens, aantallen afnemers en aansluitingen) positief gecorreleerd zijn met de Uitvoeringskosten. Om kosteninschattingen te kunnen baseren op enkel de lengte (prijs per meter) lijkt niet realistisch, daarvoor zouden we een correlatie waarde veel dichterbij 1 verwachten van lengte. Naast de omvang-gerelateerde variabelen zijn er nog andere variabelen die een correlatie hebben met deze kosten, waaronder ook een aantal omgevingseigenschappen.

De rechter tabel laat ook zien dat de project omvang (geplande duur en lengte) gecorreleerd zijn met de aanleg kosten, in dit geval per meter. We zien hier dat grotere projecten lagere kosten per meter hebben (negatieve correlatie coëfficiënt). Dit lijkt aannemelijk omdat bijvoorbeeld overheadkosten relatief kleiner zijn bij grotere projecten dan bij kleinere projecten, of omdat vertragingen opgevangen kunnen worden binnen een langere uitvoeringstijd. Het merendeel van de gecorreleerde variabelen zijn omgevingseigenschappen;

- ✓ Watergerelateerd: de aanwezigheid van een waterlichaam in het projectgebied, het toepassen van bemaling en het grondwaterniveau
- ✓ Gebouwgerelateerd: het gemiddelde bouwjaar van de gebouwen in het projectgebied, en standaarddeviatie van dat gemiddelde bouwjaar, voetafdruk van gebouwen in het projectgebied
- ✓ Weg- en beschikbare ruimte gerelateerd: gemiddelde afstand tussen de weg en bebouwing, wegbreedtes, verkeersgebied type
- ✓ Vervuiling- en grondgerelateerd: trefkans archeologie, aanwezigheid van vervuild oppervlak vervuiling en toepassing van grondverbetering,

Er zijn een aantal variabelen geel gemarkeerd in de tabellen. Dit zijn de variabelen die een correlatie aangeven met beide uitvoeringskosten variabelen (totaal en per meter) maar in alle gevallen een omgedraaid effect. De variabelen die een positieve correlatie laten zien bij uitvoeringskosten laten een negatieve correlatie zien bij de *M\_uitvoeringskosten* en andersom

### 4.3 Modelanalyse – Multivariate regressie

Multivariate regressie analyse is de meest gebruikte aanpak, voor warmtenetten en andere infrastructuur (water en gas), om kosten te voorspellen. Bij kleine en lineaire datasets werkt deze aanpak relatief beter dan andere (machine learning) modellen. Daarnaast kan dit model het makkelijkst worden geïnterpreteerd en/of gebruikt. De dataset die in deze exercitie wordt gebruikt en daarmee ook deze model analyse in deze context kan het best worden benaderd als een techniek om relaties te onderzoeken dan dat het leidt tot een statistische betrouwbaar kosten model.

Met 63 projecten en 51 features lijkt overfitting<sup>3</sup> een valkuil. Dat betekent dat er een selectie gemaakt moet worden van de mee te nemen variabelen in het model. Een vuistregel is een ratio tussen het aantal projecten en het aantal verklarende variabelen (features) van 15:1 of 20:1, in ieder geval niet kleiner dan 5:1. Voor 63 projecten betekent dit maximaal 12 verklarende variabelen, maar liever minder.

Om tot een vermindering van verklarende variabelen te komen worden twee strategieën toegepast en met elkaar vergeleken. De eerste aanpak is op basis van de resultaten uit de correlatie analyse – Dit zijn de ‘veel belovende’ variabelen waarmee het modelleer proces gestart wordt, vervolgens worden er telkens variabelen weggehaald die niet relevant zijn totdat de prestatie van het model het grootst is. De tweede strategie is de *forward stepwise* aanpak waarbij in de verschillende iteraties telkens de veel belovende variabelen worden toegevoegd totdat de prestatie van het model het grootst is. Op deze wijze worden verschillende modellen en combinaties van variabelen met elkaar vergeleken. Voor beide aanpakken wordt gebruikt gemaakt van SPSS software.

Als criterium voor de prestatie van deze modellen wordt gekeken naar de zogenoemde adjusted  $R^2$ . De  $R^2$  geeft aan hoeveel van de variantie in de afhankelijke variabele (kosten) verklaard wordt door de verklarende variabelen. De waarde ligt tussen de 0 en 1 waarbij 1 het best mogelijke model aangeeft. Bij meervoudige regressie moet gecorrigeerd worden voor het feit dat meer verklarende variabelen altijd meer van de variantie kunnen verklaren. Daarvoor is de adjusted  $R^2$ .

De analyses worden twee keer uitgevoerd voor beide afhankelijke variabelen: *Uitvoeringskosten* en *M\_uitvoeringskosten*. Alleen het best presterende model wordt weergegeven hieronder. Door de gekozen aanpak komen verschillende groepen variabelen naar voren als meest geschikte variabelen om de totale aanlegkosten respectievelijk de aanlegkosten per meter warmtetracé te voorspellen.

### 4.3.1 Model: Uitvoeringskosten

Figuur 4-41 en Figuur 4-42 geven de resultaten weer van het best presterende model (adjusted  $R^2$  van 0.862) voor de kosteninschatting van de Uitvoeringskosten. Meer informatie over dit model (rapportage prestatie ten opzicht van andere modellen en de rapportage van data aanpassingen) is te vinden in bijlage D.

Het eerste figuur geeft het gewicht aan van de verschillende variabelen in hun voorspellende vermogen. Te zien is dat de postcode het belangrijkste is, gevolgd door de lengte van het project. De overige variabelen die van belang zijn vallen onder omgevingseigenschappen (peilgebied niveau, vervuilingscategorie, voetafdruk weg, bouwjaar en woongebied ) en technisch eigenschappen (aantal aansluitingen, leiding diameter en supply temperatuur).

Figuur 4-42 geeft de coëfficiënten weer van het model. Dit betreft het model waarin de totale uitvoeringskosten worden voorspeld aan de hand van de opgenomen voorspellende variabelen. Onder de tabel is een toelichting opgenomen van variabelen die afwijken van de eerder genoemde

<sup>3</sup> Overfitting is een term, in machine learning, die gaat over het passen van de data op het algoritme. Een algoritme is “overfitted” als het leert om het gelijk te stellen met het data punt en niet met de onderliggende mechanismen van de data punten.

categorieën. Bijvoorbeeld, grondvervuilingscategorie 0 en 1 zijn samengevoegd. Ook de postcodes zijn ondergebracht in clusters.

De postcode van het aanlegproject is een belangrijke voorspeller. Postcode gebieden in groep 3 (1100, 1400, 3300, 6500) zijn de duurste gebieden. Postcodegebieden in groep 0 (3500, 3800, 4800, 5000, 5200, 6800) zijn bijna 100.000 € goedkoper dan de gebieden in groep 3. De lengte (totaal leidingen) van het aan te leggen warmtenet is de volgende belangrijke voorspeller. Met elke extra meter lengte komt er zo'n 475 € aan kosten bij. Hoe meer aansluitingen er gekoppeld zijn aan het warmtenet, hoe duurder het project, bij elke extra aansluiting worden de uitvoeringskosten met 2175 € duurder. Ook grotere leiding diameters hebben een verhogend effect op de uitvoeringskosten.

Vervolgens zijn er nog een aantal omgevingseigenschappen die belangrijke voorspellende waarde hebben voor de uitvoeringskosten; het peilgebied niveau, grondvervuilingscategorie voetafdruk weg (grotere voetafdruk van de weg is goedkoper), bouwjaar (hoe recenter hoe goedkoper), en woongebied (geen woongebied is goedkoper). Ook een toenemende supply temperatuur heeft een verhogende werking op uitvoeringskosten.

Het model kan als volgt worden gelezen

**Uitvoeringskosten van project X=**

**$5.60 \cdot 10^6$**

**$+ (0 \text{ als postcode}=3, -753 \cdot 10^3 \text{ als postcode}=2, -840 \cdot 10^3 \text{ als postcode is}=1, -950 \cdot 10^3 \text{ als postcode}=0)$**

**$+ (476 \cdot \text{lengte [m]})$**

**$+ (2,175 \cdot \text{aantal aansluitingen}[\#])$**

**$+ (940 \cdot \text{leiding diameter [mm]})$**

**$+ (46.1 \cdot 10^3 \cdot \text{peilgebiedniveau [mNAP]})$**

**$+ (65.7 \cdot 10^3 \text{ als grondvervuiling}=0, -260.7 \cdot 10^3 \text{ als grondvervuiling is } 1, 0 \text{ als grondvervuiling is } 2)$**

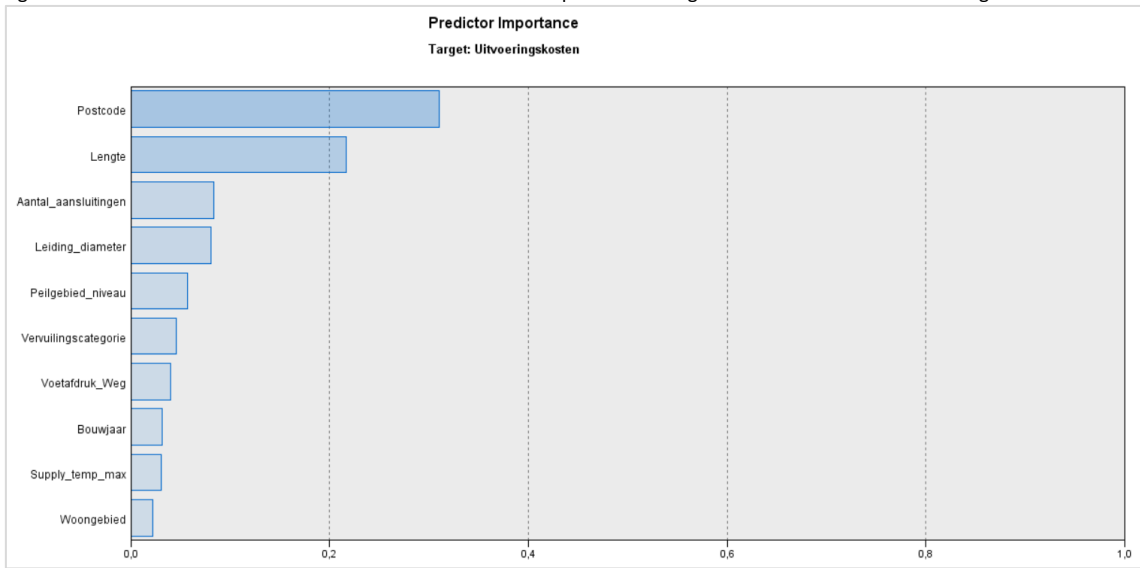
**$+ (-515 \cdot 10^3 \cdot \text{voetafdrukweg [fractie tussen 0 en 1]})$**

**$+ (-985 \cdot \text{bouwjaar[jaartal]})$**

**$+ (3.11 \cdot 10^3 \cdot \text{supplytemp}[^\circ\text{C}])$**

**$+ (0 \text{ als woongebied}=1, -78.9 \cdot 10^3 \text{ als woongebied}=0)$**

Figuur 4-41 Zwaarte van de verschillende variabelen in het voorspellend vermogen van het model voor Uitvoeringskosten





Figuur 4-42 Model coëfficiënten Uitvoeringskosten

**Coefficients**  
**Target: Uitvoeringskosten**

Model Term	Coefficient ▶	Sig.	Importance
Intercept	5.600.619,083	,010	
Postcode_transformed=0	-949.889,137	,000	0,310
Postcode_transformed=1	-840.461,940	,000	0,310
Postcode_transformed=2	-753.263,338	,000	0,310
Postcode_transformed=3	0,000 <sup>a</sup>		0,310
Lengte_transformed	475,955	,000	0,217
Aantal_aansluitingen_transformed	2.175,056	,002	0,084
Leiding_diameter_transformed	940,444	,003	0,080
Peilgebied_niveau_transformed	46.092,255	,011	0,057
Grondvervuiling_transformed=0	65.740,789	,491	0,045
Grondvervuiling_transformed=1	-260.681,776	,139	0,045
Grondvervuiling_transformed=2	0,000 <sup>a</sup>		0,045
Voetafdruk_Weg_transformed	-514.534,726	,031	0,040
Bouwjaar_transformed	-985,517	,054	0,031
Supply_temp_max_transformed	3.108,347	,058	0,030
Woongebied=0	-78.873,244	,104	0,022
Woongebied=1	0,000 <sup>a</sup>		0,022

<sup>a</sup>This coefficient is set to zero because it is redundant.

***Transformed variables:***

*postcode 0 = 3500, 3800, 4800, 5000, 5200, 6800 ; postcode 1 =*

*2500, 3400, 6200, 7500 ; postcode 2 = 1000, 2300, 2400, 3000*

*postcode 3 = 1100, 1400, 3300, 6500*

*Grondvervuiling 0 = vervuilingscategorie 0 en 1*

*Grondvervuiling 1 = vervuilingscategorie 2*

*Grondvervuiling 2 = vervuilingscategorie 3*

### 4.3.2 Model: M\_Uitvoeringskosten

Figuur 4-43 en Figuur 4-44 geven de resultaten weer van het best presterende model (adjusted R<sup>2</sup> van 0.559) voor de kosteninschatting van de per meter uitvoeringskosten. Meer informatie over dit model (rapportage prestatie ten opzicht van andere modellen en de rapportage van data aanpassingen) is te vinden in bijlage E.

Het eerste figuur geeft het gewicht aan van de verschillende variabelen in hun voorspellende vermogen. Te zien is dat de variabele waterlichaam het belangrijkste is, gevolgd door de wegbreedte en het postcode gebied. Delta temperatuur en diepte opensleuf zijn de enige technische eigenschappen die in het model zijn opgenomen, de overige zijn omgevingseigenschappen (bouwjaar, voetafdrukweg, voetafdruk bomen, voetafdruk wegtype kruising, aandeel vervuilde grond).

Figuur 4-44 geeft de coëfficiënten weer van het model. Dit betreft het model waarin de per meter uitvoeringskosten worden voorspeld aan de hand van de opgenomen voorspellende variabelen. Onder de tabel is een toelichting opgenomen van variabelen die afwijken van de eerder genoemde categorieën, in dit geval het groeperen van de postcodes.

Wanneer er geen waterlichaam aanwezig is in het project gebied kunnen de per meter uitvoeringskosten zo'n 3000 € lager zijn dan wanneer er wel een waterlichaam aanwezig is. Een bredere weg, de tweede voorspellende variabele, zorgt ook voor lagere per meter uitvoeringskosten per extra meter breedte gaan de per meter kosten met 236€ omlaag. De postcode van het aanlegproject is net als in het vorige model een belangrijke voorspeller. Postcode gebieden in groep 0 (3500, 3800, 4800, 5000, 5200, 6800 ) hebben hogere per meter uitvoeringskosten, het scheelt zo'n 1000 € per meter ten opzichte van de overige postcode gebieden uit de data set.

Hoe recenter het gemiddelde bouwjaar, hoe lager de per meter uitvoeringskosten. Dit bedrag gaat met 8 € omlaag bij het toenemen van het gemiddelde bouwjaar. Wanneer het project gebied te maken heeft met een hoge wegvoetafdruk nemen de kosten per meter flink toe, met ongeveer 45€ per 1% hogere wegvoetafdruk en 450euro bij 10% hogere wegvoetafdruk. Daarbij is het wel van belang te realiseren dat de dataset waarop dit model is gebaseerd een max wegvoetafdruk heeft van 0.6 (60%). Ook met een diepere opensleuf zijn de uitvoeringskosten hoger. Per meter diepte komt er zo'n 1000€ bij. Een hoge voetafdruk bomen heeft een prijsverlagend effect, net als de voetafdruk van wegtype kruising. Wanneer het aandeel vervuilde grond toeneemt, nemen de per meter kosten af, en wanneer de delta temperatuur toeneemt neemt ook de per meter prijs toe.

Het model kan als volgt worden gelezen (geen cijfers achter de komma weergegeven)

**Per meter Uitvoeringskosten (€) van project X =**

***20.49\*10<sup>3</sup>***

***+ (-3297 als waterlichaam niet aanwezig is, 0 als waterlichaam wel aanwezig is)***

***+ (- 236 \*wegbreedte[m])***

***+ (1,086 als postcode = 0, 0 als postcode = 1)***

***+ (-8\*bouwjaar[jaartal])***

***+ (4,515 \* voetafdruk weg[fractie tussen 0 en 1])***

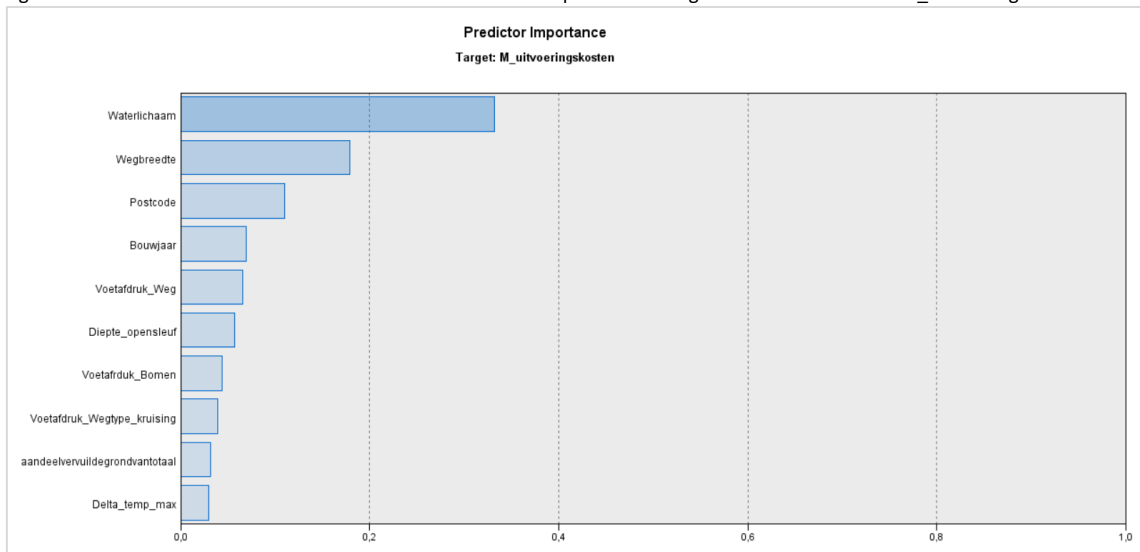
***+ (1,108 \* diepte opensleuf [m])***

***+ (-25 \*voetafdruk bomen[fractie tussen 0 en 1])***

***+ (-6,611 \* voetafdruk wegtype kruising [fractie tussen 0 en 1])***

**+ (-1,206 \*aandeel vervuilde grond[fractie tussen 0 en 1])**  
**+ (19\*delta temp [°C])**

Figuur 4-43 Zwaarte van de verschillende variabelen in het voorspellend vermogen van het model voor M\_Uitvoeringskosten



Figuur 4-44 Model coëfficiënten M\_Uitvoeringskosten

**Coefficients**  
Target: M\_uitvoeringskosten

Model Term	Coefficient ▶	Sig.	Importance
Intercept	20.493,006	,004	
Waterlichaam=0	-3.297,654	,000	0,332
Waterlichaam=1	0,000 <sup>a</sup>		0,332
Wegbreedte_transformed	-236,316	,000	0,179
Postcode_transformed=0	1.086,434	,003	0,110
Postcode_transformed=1	0,000 <sup>a</sup>		0,110
Bouwjaar_transformed	-8,092	,016	0,070
Voetafdruk_Weg_transformed	4.515,213	,019	0,065
Diepte_opensleuf_transformed	1.108,717	,029	0,057
Voetafdruk_Bomen_transformed	-25,902	,053	0,044
Voetafdruk_Wegtype_kruising_transformed	-6.611,039	,068	0,039
Aandeelvervuildegrondvantotaal_transformed	-1.206,251	,101	0,031
Delta_temp_max_transformed	19,679	,109	0,030

<sup>a</sup>This coefficient is set to zero because it is redundant.

*Transformed variables:*

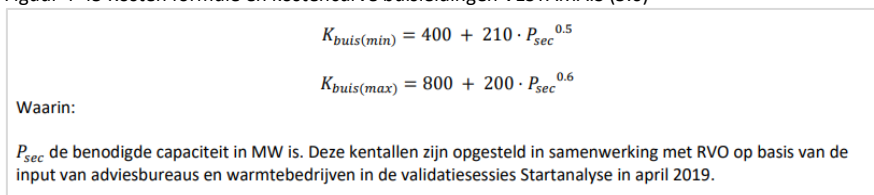
*postcode 0 = 3500, 3800, 4800, 5000, 5200, 6800 ;  
postcode 1 = 2500, 3400, 6200, 7500, 1000, 2300, 2400, 3000,  
1100, 1400, 3300, 6500*

### 4.3.3 Model: Baseline

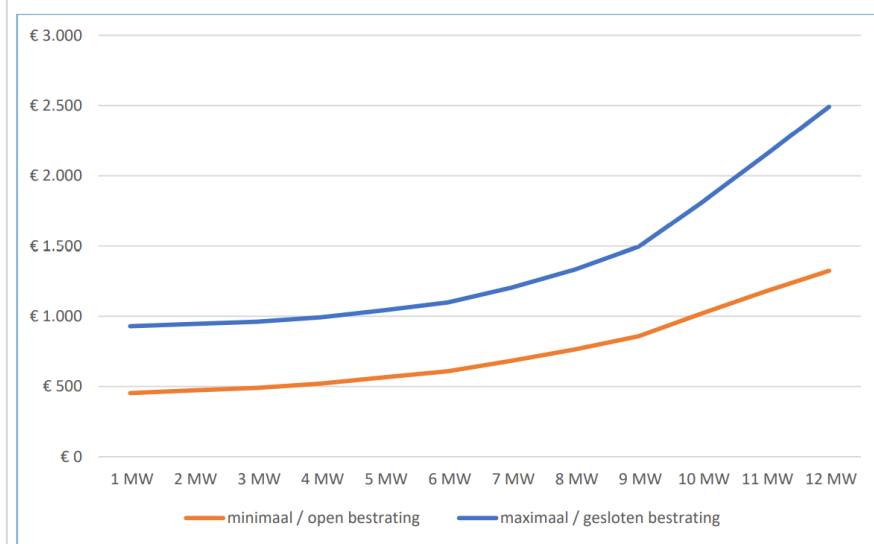
Bij het trainen en optimaliseren van een lineair (multivariate) regressiemodel, of een ander ML-model, is het belangrijk om een scorecriterium (in ons geval de adjusted Rsquare) en een minimaal haalbaar prestatieniveau te identificeren. Dit minimaal haalbare prestatieniveau, ook wel de baseline genoemd, vertegenwoordigt de momenteel beschikbare modellen. Wanneer het nieuw ontwikkelde model minder presteert dan de baseline biedt het nieuwe model geen toegevoegde waarde. Dat zou betekenen dat het model verder doorontwikkeld moet worden, of dat het model ongeschikt(er) is voor de betreffende toepassing. Het scorecriterium is noodzakelijk om de prestaties van de modellen te kunnen bestuderen en vergelijken. Met andere woorden, hoe goed kan het model op basis van de voorspellende variabelen de waarden van de response variabelen voorspellen? In de volgende modelanalyses voor het baseline model wordt ook de  $R^2$  bekeken.

Voor het opstellen van een baseline model is een pragmatische aanpak gekozen en wordt er een vergelijking getrokken met de rekenmethode zoals wordt toegelicht in het functioneel ontwerp van Vestamais (5.0). Voor het aanleggen van buisleidingen – warmteleidingen in de grond – wordt in Vesta MAIS gerekend met een bedrag per meter welke een functie is van de benodigde capaciteit (de diameter van de buis). De formule voor buisleidingen kent een bovengrens en een ondergrens welke vooral gekoppeld is aan het type bestrating. In ons baseline model wordt de bovengrens gebruikt.

Figuur 4-45 Kosten formule en kostencurve buisleidingen VESTAMAIS (5.0)



FIGUUR 7.2 – KOSTENCURVE WARMTELEIDINGEN (KOSTEN PER METER BUISLEIDING)



Om de prestatie van de in 4.3.1 en 4.3.2. gepresenteerde modellen voor Uitvoeringskosten en M\_Uitvoeringskosten te vergelijken met de een baseline model worden de volgende modellen opgezet.

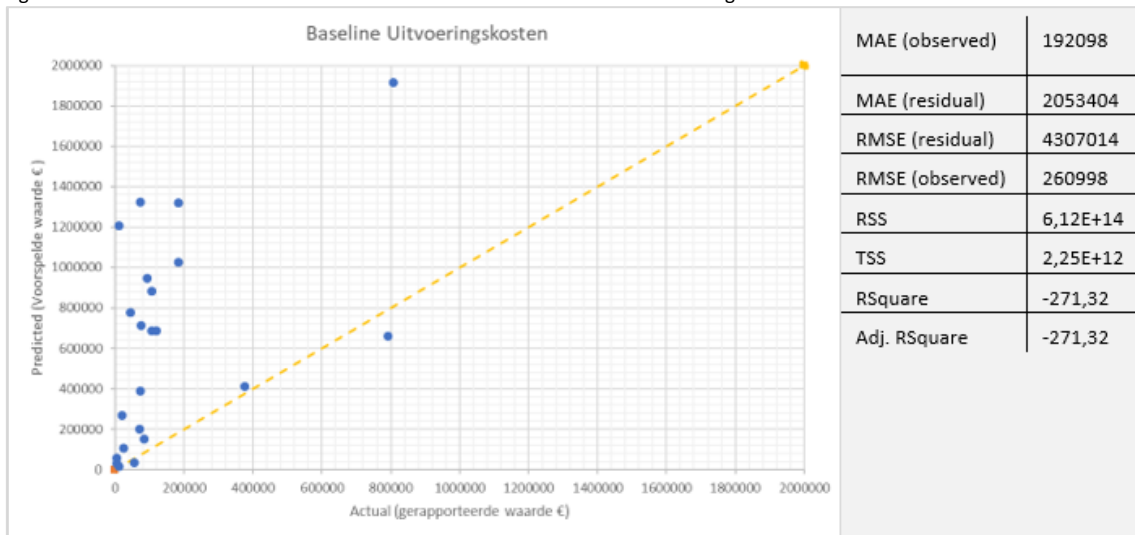
$$\text{Baseline Uitvoeringskosten} = (800 + 200 * P_{\text{sec}}^{0.6}) * \text{projectlengte}$$

$$\text{Baseline M\_Uitvoeringskosten} = (800 + 200 * P_{\text{sec}}^{0.6})$$

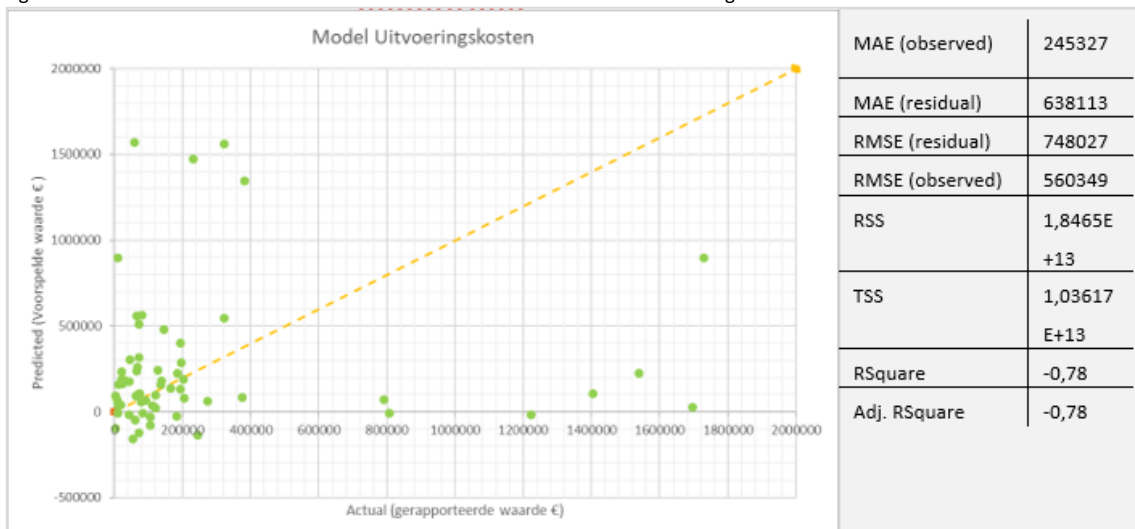
Helaas zijn niet voor alle projecten ook de aangelegde vermogens bekend. Van de 63 projecten zijn er 33 met gegevens over het vermogen waardoor deze gebruikt kunnen worden voor het baseline model. Dit bemoeilijkt de vergelijking van prestatie indicatoren.

Wanneer voor de Uitvoeringskosten het baseline model wordt vergeleken met het ontwikkelde regressie model zien we dat het regressie model wel beter presteert (lagere MAE en RMSE) maar beide modellen lijken ondermaats. Bij het regressie model is vooral te zien dat bij de goedkopere projecten, onder de 40k, de voorspelde kosten wel redelijk in de buurt liggen van de werkelijke (gerapporteerde) kosten.

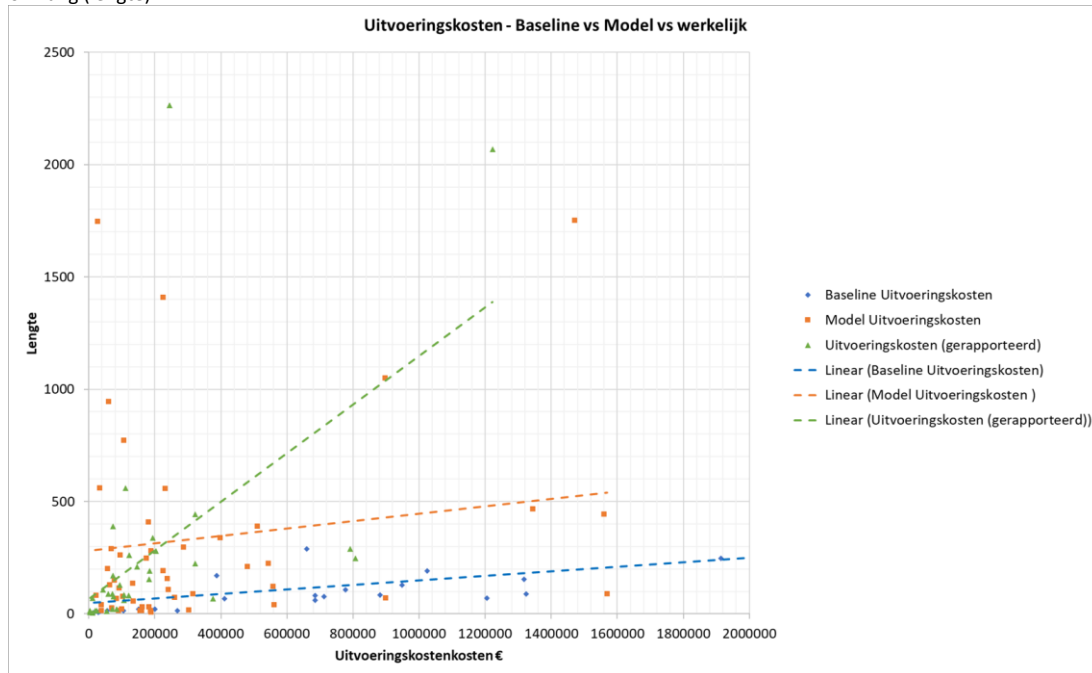
Figuur 4-46 Grafiek en evaluatie criteria Predicted vs Actual voor Baseline Uitvoeringskosten



Figuur 4-47 Grafiek en evaluatie criteria Predicted vs Actual voor Baseline Uitvoeringskosten

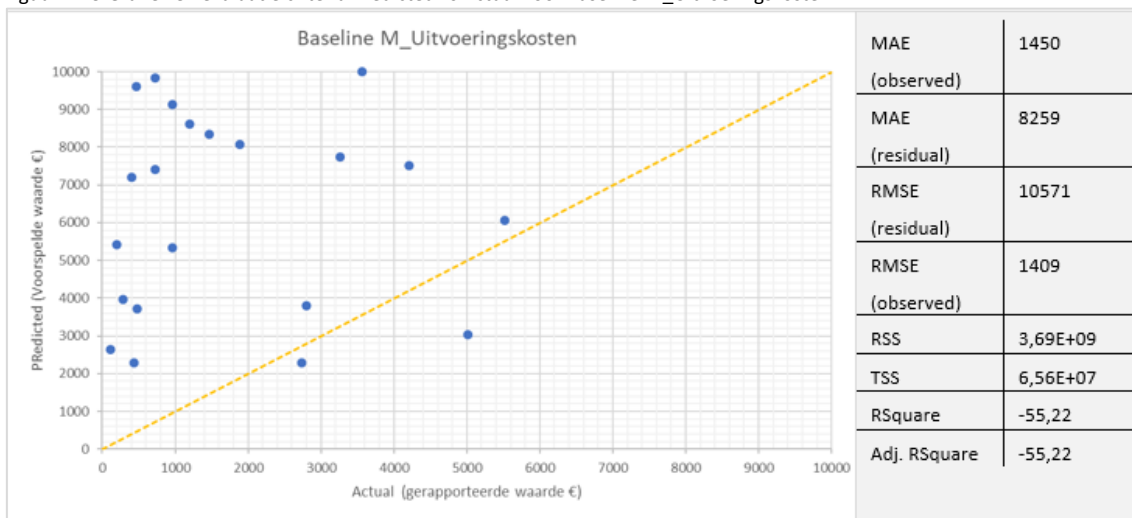


Figuur 4-48 Baseline, Regressie model en werkelijke uitvoeringskosten – trendlijn vergelijking tussen projecten van verschillende omvang (lengte)

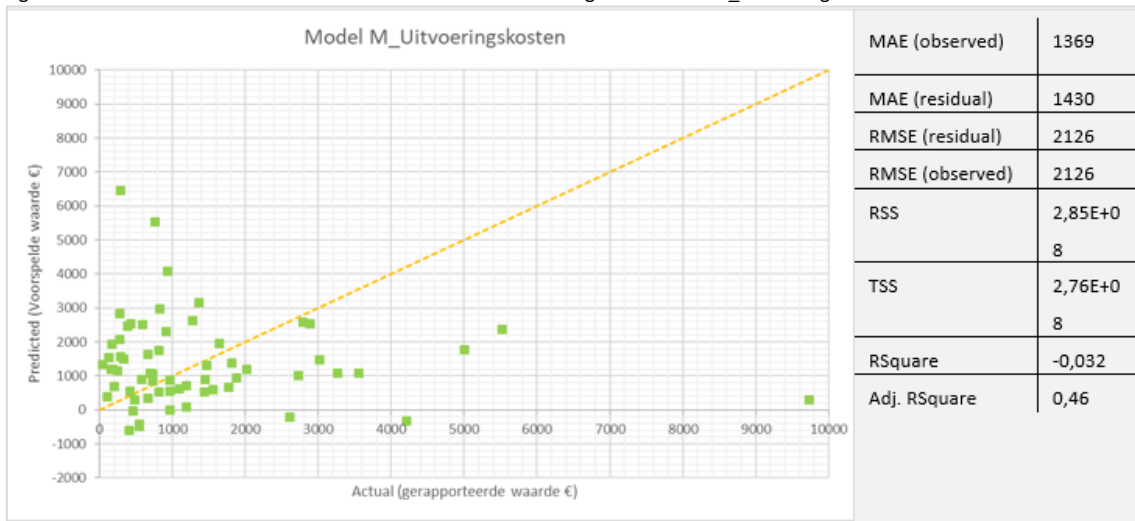


Wanneer voor de per meter uitvoeringskosten (M\_Uitvoeringskosten) het baseline model wordt vergeleken met het ontwikkelde regressie model zien we ook dat het regressie model beter presteert (lagere MAE en RMSE) en het regressie model lijkt nog een redelijke voorspellende waarde te hebben (adjusted R-Square van 0,46). Deze presentatie wijkt af van de prestatie van het beschreven model in 4.3.2 omdat er in dat model anders is omgegaan met outliers en missende waarden. In deze huidige vergelijking wordt alle data gebruikt zoals aangeleverd en beschikbaar. Wederom lijkt het regressie model beter te presteren voor de lagere uitvoeringskosten per meter.

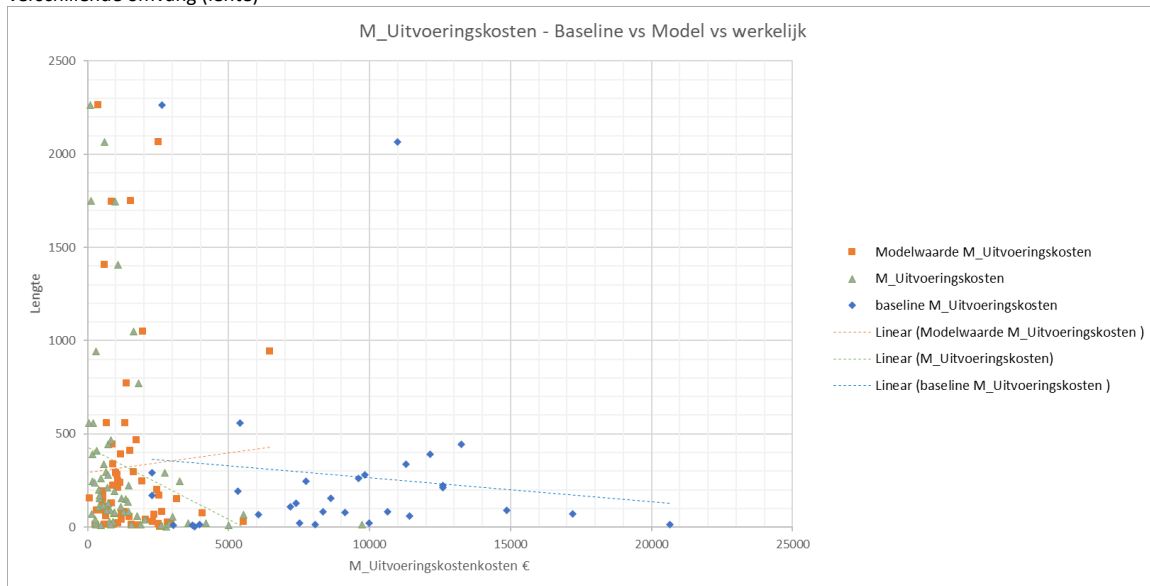
Figuur 4-49 Grafiek en evaluatie criteria Predicted vs Actual voor Baseline M\_Uitvoeringskosten



Figuur 4-50 Grafiek en evaluatie criteria Predicted vs Actual voor regressie model M\_Uitvoeringskosten



Figuur 4-51 Baseline, Regressie model en werkelijke M\_Uitvoeringskosten –trendlijn vergelijking tussen projecten van verschillende omvang (lente)





## 5 Conclusies en aanbevelingen

Het doel van dit project, is ten eerste om beter inzicht te krijgen in de factoren die in de bebouwde omgeving invloed hebben op de aanlegkosten van warmtenetten en ten tweede om deze inzichten in te zetten om te komen tot realistischere kosteninschattingen van warmtenet aanlegprojecten in de bebouwde omgeving. Project 2C beoogt daarmee bij te dragen aan het verkleinen van de bestaande bandbreedte van onzekerheid bij kosteninschattingen van warmtenet aanlegprojecten door rekening te houden met de invloed van omgevingsfactoren op de aanlegkosten. Onderliggend is daarmee de onderzoekshypothese dat omgevingsfactoren van significante invloed zijn op de aanlegkosten en dat het niet meenemen van deze omgevingsfactoren bij kosteninschattingen een verklaring is van een deel van de bestaande bandbreedte van onzekerheid.

### 5.1 Conclusies

Er is in de modelanalyse enkel gekeken naar de uitvoeringskosten voor terreinleidingen, exclusief materiaalkosten en de kosten voor inpandigte aanpassingen. Op basis van het onderzoek kunnen we de volgende conclusies trekken.

#### 5.1.1 Relevante omgevingseigenschappen

Vanuit de literatuur, aangevuld door praktijk ervaring van de projectpartners is een reeks aan omgevingseigenschappen geïdentificeerd die van invloed zijn op de aanlegkosten van een warmtenet. Dit zijn eigenschappen, naast technische eigenschappen (die veelal gebruikt worden voor kosten inschattingen) die (naar verwachting) van invloed zijn op de te verwachten uitvoeringskosten. Een aantal van de door literatuur en praktijk geïdentificeerde factoren lijken ook in de data analyse bevestigd te worden. Niet alle factoren konden onderzocht worden doordat niet alle data beschikbaar was of er onvoldoende variatie in de database was.

Tabel 5.1 Omgevingsfactoren van invloed op de Uitvoeringskosten en uitvoeringskosten per meter (kolom M\_ Uitvoeringskosten)

Geïdentificeerd in literatuur	Aangevuld door praktijk ervaring	Data Analyse (niet statistisch bewezen)	
		Uitvoeringskosten	M_ Uitvoeringskosten
Stedelijke/landelijke omgeving	Niet gesprongen explosieven	Postcode	Waterlichaam
Weg bekleding	Archeologie	Peilgebied niveau	Wegbreedte
Verkeersdrukte	Water (bemaling nodig)	Vervuilingscategorie	Postcode
Kwaliteit van de grond	Wegtype	Voetafdruk weg	Bouwjaar
Type gebouwen	Beschikbare werkruimte	Bouwjaar	Voetafdruk weg
Drukte in de ondergrond	bovengronds	Woongebied	Voetafdruk bomen
Grond vervuiling	Bomen aanwezig		Voetafdruk wegtype kruising
Teer in het asfalt			Aandeel vervuilde grond
Voetgangershellingen, trottoirs, en stoepanden			

De onderzochte database is te klein om voldoende statistische onderbouwing te leveren. Echter, bovenstaande bevindingen wijzen er wel op dat er een rol is voor omgevingsfactoren in hun bijdrage aan het voorspellen van de uitvoeringskosten van warmtenetten. Dit is ook een conclusies die door de partners wordt gedragen.

## 5.1.2 Relevante project- en technische eigenschappen

Ook zijn er project en technische eigenschappen geïdentificeerd op basis van literatuur die van invloed zijn op de aanlegkosten van een warmtenet. Het merendeel van deze factoren worden bevestigd in de data analyse. Aanlegmethode was volgens de praktijkervaring wel een relevante factor, echter, er was onvoldoende variatie in de database om dit te onderzoeken.

Tabel 5.2 project en technische factoren van invloed op de Uitvoeringskosten en uitvoeringskosten per meter (kolom M\_Uitvoeringskosten)

Geïdentificeerd in literatuur	Aangevuld door praktijk ervaring	Data Analyse (niet statistisch bewezen)	
		Uitvoeringskosten	M_Uitvoeringskosten
Warmtevraag Lengte van leidingen (per huis) Diameter van de leidingen Overdimensionering (voor toekomstig gebruik) Zandvulling en geïmporteerde grond Gebruikte isolatie en materiaal Diepte uitgraving Zijleidingen, nieuwe leidingen, verlaten leidingen	Aanlegmethode	Lengte Aantal aansluitingen Leiding diameter Supply temperatuur	Diepte opensleuf Delta temperatuur

## 5.1.3 Kostenmodel

Zoals meermaals aangegeven is de dataset klein en foutgevoelig. In andere wetenschappelijke studies lijkt dit ook het geval te zijn. De kwantiteit en conclusies moeten daarom voorzichtig geïnterpreteerd worden.

De kostenmodellen met de hoogste prestatie, hoger dan de baseline modellen zijn:

<p><b>Voor uitvoeringskosten totaal:</b></p> <p>Uitvoeringskosten van project X=</p> $5,60 * 10^6 + (0 \text{ als postcode } 3, -753 * 10^3 \text{ als postcode } 2, -840 * 10^3 \text{ als postcode is } 2, -950 * 10^3 \text{ als postcode}=0) + (477 * \text{lengte (m)}) + (2,175 * \text{aantal aansluitingen (\#)}) + (940 * \text{leiding diameter [mm]}) + (46.1 * 10^3 * \text{peilgebiedniveau (m)}) + (65.7 * 10^3 \text{ als grondvervuiling}=0, -261 * 10^3 \text{ als grondvervuiling is } 1, 0 \text{ als grondvervuiling is } 2) + (-515 * 10^3 * \text{voetafdrukweg (-)}) + (-985 * \text{bouwjaar (\#)}) + (3.11 * 10^3 * \text{supplytemp (}^\circ\text{C)}) + (0 \text{ als woongebied}=1, -78.9 * 10^3 \text{ als woongebied}=0)$ <p><i>postcode 0 = 3500, 3800, 4800, 5000, 5200, 6800 ; postcode 1 = 2500, 3400, 6200, 7500 ; postcode 2 = 1000, 2300, 2400, 3000 postcode 3 = 1100, 1400, 3300, 6500</i> <i>Grondvervuiling 0 = vervuilingscategorie 0 en 1</i> <i>Grondvervuiling 1 = vervuilingscategorie 2</i> <i>Grondvervuiling 2 = vervuilingscategorie 3</i></p>
<p><b>Voor uitvoeringskosten per meter:</b></p> <p>Per meter uitvoeringskosten van project X =</p> $20.49 * 10^3 + (0 \text{ als waterlichaam aanwezig is, } - 3,297 \text{ als waterlichaam niet aanwezig is}) + (- 236 * \text{wegbreedte (m)}) + (1,086 \text{ als postcode } = 0, 0 \text{ als postcode } = 1) + (-8 * \text{bouwjaar}) + (4,515 * \text{voetafdruk weg (-)}) + (1,108 * \text{diepte opensleuf (m)}) + (-25 * \text{voetafdruk bomen (-)}) + (-6,611 * \text{voetafdruk Wegtype kruising (-)}) + (-1,206 * \text{aandeel tracé met vervuilde grond (-)}) + (19 * \text{delta temp (}^\circ\text{C)})$ <p><i>postcode 0 = 3500, 3800, 4800, 5000, 5200, 6800 ; postcode 1 = 2500, 3400, 6200, 7500, 1000, 2300, 2400, 3000, 1100, 1400, 3300, 6500</i></p>

Deze modellen zijn vanwege de kleine data set niet statistisch voldoende onderbouwd. De orde groottes en de richting van de effecten zijn besproken in het projectteam. Over het algemeen waren de bevindingen wel verklaarbaar met een paar uitzonderingen. De huidige database heeft naast de kleine omvang ook nog het probleem dat de verschillende bedrijven dominant zijn in bepaald gebied, of juist enkel kleinere projecten hebben aangeleverd. Daardoor is er veel onderlinge correlatie tussen de variabelen bedrijf, type contract, postcode, en project grootte. Veel van de variabelen hebben te maken met de beschikbare ruimte, zowel ruimte voor de plaatsing van de leidingen, als “werkruimte” om bijvoorbeeld materieel kwijt te kunnen.

#### **5.1.4 Baseline modellen**

De baseline modellen zijn opgesteld op basis van de rekenmethode van Vesta Mais. Voor het aanleggen van buisleidingen – warmteleidingen in de grond – wordt in Vesta MAIS gerekend met een bedrag per meter welke een functie is van de benodigde capaciteit (de diameter van de buis). Deze aanpak wordt toegepast op de data set, om door middel van het gebruik van de functie de verwachte kosten te voorspellen. Deze wijkt sterk af van de gerapporteerde (werkelijke) waarde van de uitvoeringskosten. De ontwikkelde regressie modellen presteren beter dan de baseline modellen, maar alsnog ondermaats (bij uitvoeringskosten) of laag (bij per meter uitvoeringskosten).

Los van de prestatie van de ontwikkelde regressie modellen lijkt de functie die gebruikt is in de baseline modellen niet goed de kosten te voorspellen voor de projecten in onze data base.

## **5.2 Aanbevelingen**

### **5.2.1 Postcode en vergunningsvoorwaarden**

Uit de data analyse blijkt dat het postcode gebied waarin een project plaatsvind veel invloed heeft op de uitvoeringskosten (zowel totaal als per meter). Mogelijk zijn er in deze variabelen onderliggende correlaties (met bijvoorbeeld project grootte) maar door de projectpartners is aangegeven dat verschillende gemeenten ook verschillende richtlijnen en vergunningsvoorwaarden aanhouden die van invloed zijn op de kosten. Dit effect lijkt ook in de modelresultaten zichtbaar.

Het is aan te raden nader uit te zoeken wat de verschillen zijn, wat dit betekent voor de aanlegkosten van warmtenetten en hoe dit zich verhoudt tegen de noodzaak van de voorwaarden en richtlijnen.

### **5.2.2 Valideren resultaten data analyse in de praktijk**

De model resultaten laten vooralsnog niet volledige statistisch betrouwbare resultaten zien, echter geeft het al wel inzicht in mogelijk relevant factoren die van invloed zijn. Op basis van het huidige model kan er door de projectpartners, maar ook door andere marktpartijen worden gerekend aan kosteninschattingen, gebruikmakend van deze resultaten. Op basis van deze ervaringen kan het model, en ook de data behoefte verbeterd worden, of kunnen er nieuwe, gedragen, kentallen voor omgevingseigenschappen tot stand komen.

### **5.2.3 Valideren resultaten data analyse met grotere database**

De data base die gebruikt is voor dit onderzoek is opgesteld ten behoeve van deze analyse. Door het projectteam wordt ervaren hoe waardevol het is om over kwantitatieve data te beschikken en als basis te gebruiken voor discussies. Er wordt dan ook door het projectteam aanbevolen om de huidige database te bewaren en gestructureerd uit te bereiden. Een grotere omvang van de beschikbare database kan leiden tot waardevolle inzichten op gebied van kosteninschattingen

maar kan daarmee ook richting geven aan de richting van (monitoren van) kostenbesparende innovaties. Daarnaast is het wenselijk om een grotere diversiteit aan postcodes, bedrijven en groottes (in lengte) van projecten in de data set op te nemen.

In het huidige onderzoek is er geëxperimenteerd met verschillende soorten modellen, waarvan het regressiemodel het beste uit de verf kwam, maar nog steeds niet overtuigend presteert. Mogelijk dat er bij grotere betrouwbaardere datasets andere modellen geschikter zijn. Daarnaast is in de huidige analyse geen rekening gehouden met inflatie effecten op de prijs. De verwachting was dat het effect van eventuele prijsfluctuaties als gevolg van inflatie minimaal zou zijn, zeker gezien de relatief korte tijdsperiode.

Er kan een afweging worden gemaakt aan de voorkant hoeveel projecten er per jaar zouden kunnen worden toegevoegd en welke omvang de database nodig zal hebben voor statistische betrouwbaardere modellen.

#### **5.2.4 Synergie met energie transitie rekenmodellen (vervolg)**

Er zijn veel verschillende energie transitie rekenmodellen, en daarbij ook verschillende rekenmethodes. De resultaten van dit huidige onderzoek tonen aan dat er verbetering mogelijk is door het meenemen van omgevingsfactoren en de kosten inschattingen. Naast de kennis die in het kader van dit project is opgedaan en aanvullende kennis die bij de marktpartijen aanwezig is, is er al veel kennis aanwezig bij de ontwikkelaars van de energie transitie rekenmodellen. Idealiter wordt op basis van de beschikbare kennis een verbeterslag geslagen in het vroegtijdig voorspellen van de verwachte aanlegkosten van warmtenetten door de rekenmodellen te updaten. Hiervoor wordt een samenwerking met de verschillende partijen sterk aanbevolen. Door de projectpartners is aangegeven dat betere inschattingen aan de voorkant zullen leiden tot een betere projectstart doordat de verschillende verwachtingen van de betrokken partijen wat betreft aanlegkosten meer op een lijn zitten en al beter onderbouwd kunnen worden.

Een aantal van de variabelen die in de kostenmodellen zitten (zoals bijvoorbeeld het bouwjaar) zijn al variabelen die reeds gebruikt worden in het Vesta MAIS model, dan wel in een andere deel van het model (niet specifiek de aanleg kosten voor terreinleidingen). De verwachting is dat deze variabelen makkelijker gekoppeld kunnen worden aan de huidige rekenmodellen omdat er geen extra data inventarisatie voor nodig is. Ook kan er worden gedacht aan het creëren van een GIS kaartlaag met de verwachte prijs per meter gekoppeld aan een specifiek gebied omdat de onderliggende relevante omgevingsfactoren daarin zijn opgenomen. Een dergelijke kaartlaag kan ook interessant zijn voor ontwerp optimalisatie tools (de kortste route hoeft niet altijd de goedkoopste te zijn).

#### **5.2.5 Aandachtpunten voor data inventarisatie**

Er zijn nog een aantal specifiek aandachtpunten waar rekening mee gehouden kan worden wanneer er in de toekomst opnieuw of aanvullende data geïnventariseerd wordt. De in dit project opgestelde template kan worden opgevraagd.

Het verzamelen van de data is gevoelig voor foute interpretatie door de “invuller”. Het aanstellen van coördinatoren/contactpersonen kan helpen. Daarnaast kan er bij verder gebruik van de template nader toelichting worden gegeven over de volgende zaken:

- De verschillen in “tracé” en de leiding lengte. Leiding lengte kan immers slaan op 2 keer de tracé lengte in geval van heen en weer.

- Eenduidig invullen van de kostenposten met of zonder opslag voor winst, en met of zonder transport en stallingskosten.
- In sommige gevallen wordt herbetegeling van straten en trottoirs niet door dezelfde partij uitgevoerd en is het van belang om aan te geven of de kosten inclusief of exclusief deze werkzaamheden zijn.
- Materiaalkosten zijn in het huidige onderzoek niet bekeken vanwege gebrek aan data. Het apart aanleveren van materiaal kosten is wel waardevol.

Bovenstaande zaken hebben in de huidige dataset mogelijk bijgedragen aan de foutgevoeligheid van de dataset.

### 5.2.6 Ontbrekende GIS data

De drukte in de ondergrond is een van de factoren die van belang lijkt te zijn voor het inschatten van de kosten voor het aanleggen van warmtenetten. In het huidige project is gebleken dat er geen openbare kaartlaag is waarin deze drukte wordt weergegeven. Er is een poging gedaan om gebruik te maken van de data van het kadaster. Echter, deze data was niet gratis openbaar toegankelijk, alleen verkrijgbaar voor specifieke tracés en geen grotere gebieden, en er bleek geen historische gegevens bijgehouden. Daarmee wordt bedoeld dat enkel de huidige situatie opgevraagd kan worden en niet de situatie van 5 jaar geleden ten tijde van de aanleg.

De aannemers en marktpartijen die een warmtenet gaan aanleggen zijn verplicht om een melding te maken, waardoor er ook bekend wordt welke andere infrastructuur aanwezig is op de projectlocatie. Echter na afronding wordt deze informatie niet meer bewaard. Het wordt aanbevolen om dat wel te doen.

Daarnaast is er geen openbare database gevonden ten behoeve van dit project van de verschillende soorten straat- en wegbekledingen. Enkel het onderscheid tussen verhard en onverhard, maar niet het type verharding. Het wordt aangeraden dat marktpartijen en aannemers dergelijke gegevens ook zelf bewaren bij project administratie.

# Bibliography

WarmingUp. (2020). *Definities warmtebranche*.

Curran, R. A., Raghunathan, S., & Price, M. (2004). *Review of aerospace engineering cost modelling : The genetic causal approach*. 40, 487–534. <https://doi.org/10.1016/j.paerosci.2004.10.001>

Dai, J. S. (2016). *Product Cost Estimation : Technique Classification*. 128(May 2006), 563–575. <https://doi.org/10.1115/1.2137750>

Duran, O., Rodriguez, N., & Airton, L. (2008). *Neural Networks for Cost Estimation of Shell and Tube Heat Exchangers. II*, 19–21.

Hueber, C., Horejsi, K., & Schledjewski, R. (2016). *Review of cost estimation : methods and models for aerospace composite manufacturing. January*. <https://doi.org/10.1080/20550340.2016.1154642>

Mieras, J. (2021). *Effect of the surrounding on the construction cost of district heating networks and similar infrastructures*. TUDelft.

Reidhav, C. Werner, S. *Investment models for district heating in areas with detached houses. 10th International Symposium on District Heating and Cooling, (September), 2006.*

Advanced: Heat infrastructure costs. <https://docs.energytransitionmodel.com/main/heat-Infrastructure-costs>. Accessed: 2020-05-28.

Overmorgen. Het Warmtetransitiemodel. Technical report, Overmorgen, Amersfoort, 2020.

Ed Kerckhoffs. Werking Caldomus Werking Caldomus. Technical report, Innoforte, Druten, 2017. URL <https://www.innoforte.nl/in/wp-content/uploads/2018/02/171209-werking-Caldomus.pdf>.




Benno Schepers, Ruud van den Wijngaart, Alexander Oei, and Maarten Hilferink. Functioneel ontwerp Vesta 4.0. Technical report, CE Delft, Delft, 2019. URL [https://www.pbl.nl/sites/default/files/downloads/pbl-2019-ce-delft-functioneel-ontwerp-vesta-4.0\\_4085.pdf](https://www.pbl.nl/sites/default/files/downloads/pbl-2019-ce-delft-functioneel-ontwerp-vesta-4.0_4085.pdf).




Thermos help page. <https://tool.thermos-project.eu/help/index.html>, . Accessed: 2020-07-02

# Bijlagen

## A. Longlist Variabelen - inhoud en bronnen

Voor alle variabelen in de database wordt in onderstaande tabellen een omschrijving gegeven van de variabele, aangegeven van welke bron deze afkomstig is, en het type variabele

<b>Nominal</b> → 	naming scale, where variables are simply “named” or labeled, with no specific order
<b>Ordinal</b> → 	variables in a specific order, beyond just naming them
<b>Scale</b> → 	a variable measurement scale that not only produces the order of variables but also makes the difference between variables known along with information on the value of true zero.




 = *Nominaal*,  = *Ordinaal*,  = *Schaal*

\*symbolen afkomstig uit SPSS

De grijze rijen onderaan elke tabel zijn de variabelen die tijdens het opschonen zijn “afgevallen”









- omdat er te weinig datapunten voor beschikbaar zijn, of
- omdat deze variabelen geen voorspellend karakter hebben. Met andere woorden, deze zijn enkel “post – aanleg” bekend. Voorbeeld is de werkelijke duur. Deze variabele kan niet gebruikt worden voor het voorspellen van de aanlegkosten omdat dit vooraf niet bekend is, geplande duur wel.
- Een of meerdere variabelen zijn gebruikt voor het creëren van een nieuwe variabele (bijvoorbeeld startjaar en startmaand vormen gezamenlijk de nieuwe variabele “tjdstempel”)

### Projecteigenschappen









Variabele Naam	Omschrijving	Bron	Type
Bedrijf	Het warmtebedrijf dat het project gerealiseerd (en de data heeft aangeleverd). <i>5 bedrijven, 5 categorieën</i>	Warmtebedrijf via template	Nominal 
Type	Soort aanleg project <i>1 = Vervanging bestaand warmtenetwerk</i> <i>2 = Aanleg inbreidingslocatie</i> <i>3 = Aansluiting bestaande bouw</i>	Warmtebedrijf via template	Nominal 
Contract	Welke werkzaamheden waren gecontracteerd? <i>1 = Uitvoering</i> <i>2 = Engineering en uitvoering</i> <i>3 = Ontwerp, Engineering en uitvoering</i>	Warmtebedrijf via template	Nominal 

Geplande duur	Hoe lang het aanlegproject naar verwachting zou gaan duren, in dagen.	Warmtebedrijf via template	Scale 
Tijdstempel	Berekend op basis van de startdatum van het project, geeft de startdatum aan op een doorlopende tijdschaal. Het meest recente project uit de database is gestart in aug 2020 en het oudste project is gestart in januari 2015	Warmtebedrijf via template	Scale 
Werkelijke duur	Hoe lang het aanlegproject daadwerkelijk heeft geduurd, in dagen.	Warmtebedrijf via template	Scale 
Meekoppelen	Tijdens de aanleg van het warmtenet zijn er gelijktijdig andere werkzaamheden uitgevoerd (bijvoorbeeld het vervangen van andere infra)? Ja = 1, Nee= 2	Warmtebedrijf via template	Nominal 
Startjaar	Jaar waarin de uitvoering van het project is gestart. Als input gebruikt voor tijdstempel	Warmtebedrijf via template	Scale 
Startmaand	Maand waarin de uitvoering van het project is gestart. Als input gebruikt voor tijdstempel	Warmtebedrijf via template	Ordinal 

## Technische eigenschappen




Variabel Naam	Omschrijving	Bron	Type
Supply temp max [°C]	De hoogst opgegeven supply temperatuur (deze was opgevraagd voor verschillende delen van het netwerk)	Warmtebedrijf via template	Scale 
Delta temp max [°C]	Her verschil tussen supply en retour temperatuur, hoogste waarde van de verschillende delen van het netwerk,	Warmtebedrijf via template	Scale 
Lengte [m]	Totale lengte van het tracé (aangelegde warmtenet)	Warmtebedrijf via template	Scale 
Aantal afnemers	Het aantal afnemers dat is aangesloten in dit project. (bij een flatgebouw gaat het dus om het aantal aangesloten appartementen) .	Warmtebedrijf via template	Scale 
Aantal aansluitingen	Het aantal aansluitleidingen naar het distributienet (bij een flatgebouw is dit dus 1 aansluiting)	Warmtebedrijf via template	Scale 
Aantal Bronnen	Aantal aan het warmtenet aangesloten bronnen (bij kleine projecten 0 als bron buiten de scope is)	Warmtebedrijf via template	Scale 
Aangesloten vermogen [kW]	Max vermogen dat is aangesloten op het warmtenet – geeft samen met het max aangelegd vermogen een beeld van de toekomstbestendigheid	Warmtebedrijf via template	Scale 
Aangelegd vermogen [kW]	Max vermogen waar het warmtenet op gedimensioneerd is. Als dit hetzelfde is als het aangesloten vermogen kunnen we geen nieuwe aansluitingen meer bij op het net.	Warmtebedrijf via template	Scale 




Warmteoverdrachtstation vermogen [kW]	Piekvermogen van het station in de projectzone	Warmtebedrijf via template	Scale 
Warmteoverdrachtstation - Aantal	Aantal stations in de projectzone	Warmtebedrijf via template	Scale 
Dominant leiding type	Meest voorkomende leidingtype in de projectzone. ( <i>Distributieleiding=1, Aansluitleiding = 2, Transportleiding = 3</i> )	Warmtebedrijf via template	Nominal 
Dominant leiding materiaal	Meest voorkomende leidingmateriaal in de projectzone ( <i>Staal-PUR-PE-VPS=1, Steelflex - flexibel staal =2, Calpex PEX_PUR_PE = 3, Overig = 4</i> )	Warmtebedrijf via template	Nominal 
Leiding gemiddelde diameter [mm]	Gemiddelde diameter van de leidingen in de projectzone	Warmtebedrijf via template	Scale 
Aanleg methode: Opensleuf	Opensleuf is gebruikt als aanlegmethode voor de leidingen in de projectzone	Warmtebedrijf via template	Nominal 
Diepte Opensleuf [m]	De diepte van de sleuf.	Warmtebedrijf via template	Scale 
Aanleg methode: Bovengronds	Er is bovengronds gewerkt voor het aanleggen van de leidingen in de projectzone	Warmtebedrijf via template	Nominal 
Supply temp distributie [°C]	<i>Gebruikt voor variabel: supply temp max</i>	Warmtebedrijf via template	Scale 
Supply temp afleverset [°C]	<i>Gebruikt voor variabel: supply temp max</i>	Warmtebedrijf via template	Scale 
Supply temp transport [°C]	<i>Gebruikt voor variabel: supply temp max</i>	Warmtebedrijf via template	Scale 
Delta temp distributie [°C]	<i>Gebruikt voor variabel: delta temp max</i>	Warmtebedrijf via template	Scale 
Delta temp afleverset [°C]	<i>Gebruikt voor variabel: delta temp max</i>	Warmtebedrijf via template	Scale 
Delta temp transport [°C]	<i>Gebruikt voor variabel: delta temp max</i>	Warmtebedrijf via template	Scale 
Mengregelstation Vermogen [kW]	Piekvermogen van het mengregel station in de projectzone	Warmtebedrijf via template	Scale 
Mengregelstation_aantal (stuks)	Aantal mengregelstations in de projectzone	Warmtebedrijf via template	Scale 
Boosterpompstation vermogen [kW]	Piekvermogen van het boosterpompstation in de projectzone	Warmtebedrijf via template	Scale 
Boosterpompstation_aantal (stuks)	Aantal boosterpompstations in de projectzone	Warmtebedrijf via template	Scale 
Dominant leiding materiaal - overig	<i>Length Staal_In_Staal_SIS (n=0) Length Flexwell_flexibel_RVS Length Casaflex_flexibel_RVS Length OCI_flexibel_RVS Length FlexalenPB_Polyolefine_foam_HDPE</i>	Warmtebedrijf via template	Nominal 
Aanleg methode - overig:	<i>Boringen N=1 Zinker 0 Persingen 0 in lucht Schacht funderingen</i>	Warmtebedrijf via template	Nominal 

## Omgevingseigenschappen

Variabel Naam	Omschrijving	Bron	Type
Postcode	Postcode gebied (2 cijfers) waar het project is uitgevoerd.	Postcode database Nederland	Nominal
Grondvervuiling categorie	De intensiteit dan de grondvervuiling aanwezig in het projectgebied. geen=0 licht=1 middel=2 zwaar=3	Warmtebedrijf via template	Ordinal
Vervuilde tracé lengte [m]	De oppervlakte in lengte van het tracé dat te maken had met grondvervuiling	Warmtebedrijf via template	Scale
Vervuild trace (fractie van totale tracé lengte [-])	Fractie van het projectgebied dat te maken had me vervuild trace. Berekend op basis van vervuilde tracélengte en totale projectlengte	Warmtebedrijf via template	Scale
Grondverbetering toegepast	Is er tijdens de uitvoeringswerkzaamheden grondverbetering toegepast? Ja/Nee	Warmtebedrijf via template	Nominal
Grondwaterniveau [mNAP]	De hoogte van het grondwater in de projectzone	Grondwaterstand.gdb via DINO LOKET	Scale
Bemaling toegepast	Is er tijdens de uitvoeringswerkzaamheden bemaling toegepast? Ja/Nee	Warmtebedrijf via template	Nominal
Wegtype: Straat	Het wegtype 'straat' is aanwezig in de projectzone	BRT	Nominal
Wegtype: hoofdweg	Het wegtype 'hoofdweg' is aanwezig in de projectzone	BRT	Nominal
Wegtype: regionale weg	Het wegtype 'regionale weg' is aanwezig in de projectzone	BRT	Nominal
Wegtype: lokale weg	Het wegtype 'lokale weg' is aanwezig in de projectzone	BRT	Nominal
Wegtype: verbinding [fractie]	fractie oppervlakte van het projectgebied met het wegtype 'verbinding' (tussen 0 en 1)	BRT	Scale
Wegtype: kruising [fractie]	fractie oppervlakte van het projectgebied met het wegtype 'kruising' (tussen 0 en 1)	BRT	Scale
Wegtype: overig verkeersgebied [fractie]	fractie oppervlakte van het projectgebied met het wegtype 'overig verkeersgebied' (tussen 0 en 1)	BRT	Scale
Wegbreedte (m)	De gemiddelde breedte van de weg in het projectgebied	BRT	Scale
Wegbreedte: >7m	Een weg breder dan 7m is aanwezig in de projectzone	BRT	Nominal
Wegbreedte: 4-7m	Een weg met een breedte tussen de 4 en 7m is aanwezig in de projectzone	BRT	Nominal
Wegbreedte: 2-4m	Een weg met een breedte tussen de 2 en 4m is aanwezig in de projectzone	BRT	Nominal
Wegbreedte: <2m	Een weg smaller dan 2m is aanwezig in de projectzone	BRT	Nominal
Wegvoetafdruk [fractie]	fractie oppervlakte van de projectzone met weg. (tussen 0 en 1)	BRT	Scale

	Weg is gebaad gedeelte voor het wegverkeer en kan meerdere types bevatten (zie wegtype)		
Gebouwoetafdruk (fractie)	fractie oppervlakte van de projectzone met gebouw (tussen 0 en 1)	BAG	Scale 
Bouwjaar (gemiddeld)	Het gemiddelde bouwjaar van de gebouwen in de projectzone	BAG	Ordinal 
Standaard deviatie Bouwjaar	De standaard deviatie van het gemiddelde bouwjaar van de gebouwen in de projectzone	BAG	Scale 
Gemiddelde Afstand tussen gebouw en weg (m)	De gemiddelde afstand tussen gebouw en weg in de projectzone - berekend	BAG	Scale 
Gemiddelde inwonersdichtheid [#/hectare]	Gemiddelde inwonersdichtheid in de projectzone	CBS	Scale 
Woongebied	Projectzone valt onder het type 'woongebied'	Woningtypering.gdb	Nominal 
Spoorlijn	Een spoorlijn is aanwezig in de projectzone	BBG	Nominal 
Waterlichaam	Een oppervlakte waterlichaam (sloot, meer, rivier,...) is aanwezig in de projectzone	BBG	Nominal 
Peilgebied niveau (m)	Het peil van het peilgebied waar de projectzone in valt .	polderpeilen	Scale 
Bomenbedekking (fractie)	fractie van de project zone bedekt met bomen (>2,5 m)	ANK 20200629_gm_Bomenkaart_v2	Scale 
Trefkans archeologie	de trefkans (hoog, middelhoog, laag) op archeologische resten.	Archeologie	Nominal 
<i>Maatregelen NGE</i>	<i>Maatregelen getroffen tijdens de uitvoering en/of voorzorgsmaatregelen omdat er mogelijk NGE aanwezig waren in de projectzone.</i>	Warmtebedrijf via template	Nominal 
<i>Maatregelen Archeologie (Ja = 1, Nee= 2)</i>	<i>Maatregelen getroffen tijdens de uitvoering en/of voorzorgsmaatregelen omdat er mogelijk archeologische vondsten aanwezig waren in de projectzone (</i>	Warmtebedrijf via template	Nominal 
<i>Dominantwoningtype_Vrijstaand</i>	<i>Vrijstaande woning is het meest voorkomende woningtype in de projectzone</i>	Woningtypering.gdb	Nominal 
<i>Dominantwoningtype_twee-onder-een-kap</i>	<i>Twee-onder-een-kap woning is het meest voorkomende woningtype in de projectzone</i>	Woningtypering.gdb	Nominal 
<i>Dominantwoningtype_overig</i>	<i>'Overige' woning is het meest voorkomende woningtype in de projectzone</i>	Woningtypering.gdb	Nominal 
<i>Bekleding_Verhard</i>	<i>De lengte van de weg in de projectzone bekleed met het type 'verhard' - Een weg met een egale verharding (asfalt, beton, klinkers, tegels, keien, etc.).</i>	BGT	Nominal 
<i>Bekleding_half_verhard</i>	<i>De lengte van de weg in de projectzone bekleed met het type 'half-verhard' - Een weg waarop de verharding slechts gedeeltelijk is aangebracht (als twee banen met klinkers, beton, tegels,etc.) of het rijvlak is verbeterd met grind of sintels</i>	BGT	Nominal 

Bekleding_onverhard	De lengte van de weg in de projectzone bekleed met het type 'onverhard' - Een weg zonder enige vorm van kunstmatige verharding, dan wel een weg hier en daar opgevuld met puin.	BGT	Nominal 
---------------------	---	-----	---

## Toelichting Woning Typeringen BAG






### Verschillende woningtypes











We onderscheiden verschillende woningtypes, namelijk:

- **Tussen- of geschakelde woning (T)**, woning die met meerdere woningen is verbonden en die daarmee een niet-repeterende dakconstructie heeft.
- **Eind- of hoekwoning (H)**, laatste van een serie rijtjeswoningen die aan één zijde is verbonden met een ander pand met adres.
- **Twee-onder-één-kap (K)**, woning die met een andere woning (niet zijnde een geschakelde woning) is verbonden en die daarmee eenzelfde doorlopende dakconstructie heeft.
- **Vrijstaande woning (V)**, woning waarvan het pand met verblijfsobject<sup>1</sup> niet direct met een ander pand met verblijfsobject is verbonden.
- **Appartement (A)**, waar twee of meer verblijfsobjecten voorkomen binnen een pand.
- **Niet wonen (N)**, het verblijfsobject wordt niet gebruikt als woning.



## Aanlegkosten

Variabel Naam	Omschrijving	Bron	Type
Aanlegkosten leidingeninfra (totaal)	De kosten gemaakt voor het aanleggen van de leidinginfrastructuur. Zie ook figuur hieronder 	Warmtebedrijf via template	Scale 
M_Aanlegkosten leidingeninfra [€/m]	De per meter kosten voor het aanleggen van de leidinginfrastructuur. Berekend op basis van de totaal aanlegkosten leidinginfr en de tracé lengte van het project.	Warmtebedrijf via template	Scale 
Materiaalkosten (totaal) [€]	De materiaalkosten voor de leidinginfrastructuur. Deze kosten zijn inclusief opslag voor bijvoorbeeld risico en winst, en bijbehorende levering en transportkosten. Zie ook figuur hieronder 	Warmtebedrijf via template	Scale 

M_Materiaalkosten [€/m]	De per meter materiaalkosten voor de leidinginfrastructuur. Berekend op basis van de totaal materiaalkosten en de tracé lengte van het project	Warmtebedrijf via template	Scale 
Overige kosten (totaal) [€]	De kosten die gemaakt zijn voor de aanleg van de leidinginfrastructuur die niet materiaalkosten zijn. Hier vallen materieel, personeel, en installatie kosten die gemaakt zijn voor de aanleg van leidinginfrastructuur onder. Ook vallen hier kosten zoals management en communicatie onder die voor het hele project zijn gemaakt. Zie figuur hieronder 	Warmtebedrijf via template	Scale 
Overige kosten (per meter) [€/m]	De per meter overige kosten voor de leidinginfrastructuur. Berekend op basis van de totaal overige kosten en de tracé lengte van het project	Warmtebedrijf via template	Scale 
<i>Ontwerpkosten</i>	<i>De kosten die gemaakt zijn voor het basis ontwerp, definitief ontwerp en detail ontwerp. De ontwerp kosten vallen buiten de scope van deze kwantitatieve analyse maar worden opgevraagd om er zeker van te zijn dat ze niet onderdeel zijn van "Overige kosten".</i>	<i>Warmtebedrijf via template</i>	<i>Scale </i>
<i>Afleversets aanschafkosten</i>	<i>De kosten die gemaakt zijn voor de aanschaf van afleversets. Deze kosten vallen buiten de scope van deze kwantitatieve analyse maar worden opgevraagd om er zeker van te zijn dat ze niet onderdeel zijn van "Materiaalkosten".</i>	<i>Warmtebedrijf via template</i>	<i>Scale </i>
<i>Afleversets installatiekosten</i>	<i>De kosten die gemaakt zijn voor de installatie van de afleversets, waaronder personeel en materieel. Deze kosten vallen buiten de scope van deze kwantitatieve analyse maar worden opgevraagd om er zeker van te zijn dat ze niet onderdeel zijn van "Overige kosten".</i>	<i>Warmtebedrijf via template</i>	<i>Scale </i>
<i>Stations materiaalkosten</i>	<i>De kosten die gemaakt zijn voor de aanschaf van materiaal gerelateerd aan de stations. Deze kosten vallen buiten de scope van deze kwantitatieve analyse maar worden opgevraagd om er zeker van te zijn dat ze niet onderdeel zijn van "Materiaalkosten".</i>	<i>Warmtebedrijf via template</i>	<i>Scale </i>
<i>Stations installatiekosten</i>	<i>De kosten die gemaakt zijn voor de installatie van de stations, waaronder personeel en materieel. Deze kosten vallen buiten de scope van deze kwantitatieve analyse maar worden opgevraagd om er zeker van te zijn dat ze niet onderdeel zijn van "Overige kosten".</i>	<i>Warmtebedrijf via template</i>	<i>Scale </i>
<i>Stations Grondaankoop</i>	<i>De kosten die gemaakt zijn voor de aanschaf van grond gerelateerd aan de plaatsing van stations. Deze kosten vallen buiten de scope van deze</i>	<i>Warmtebedrijf via template</i>	<i>Scale </i>

	<i>kwantitatieve analyse maar worden opgevraagd om er zeker van te zijn dat ze niet onderdeel zijn van "Overige kosten".</i>		
--	--	--	--

## B. Template data inventarisatie warmtebedrijven



### Algemene gegevens

Datum

#### Contactgegevens

Contactpersoon	Email	Telefoonnummer

#### Organisatie

ID	Naam	Rol
#N/A		

#### Project en contract gegevens

Type infrastructuur	Type aanleg/vervangng	Gecontracteerde werkzaamheden
Warmtenet		

#### Beschrijving

Geef een bondige beschrijving van de werkzaamheden en scope van het project



### Onderverdeling detail niveaus van het netwerk

Zijn de kosten gegevens (deels) beschikbaar op gedetailleerder niveau? Hier worden die niveaus gespecificeerd.

#### Netwerkgegevens hoofdproject

Deelproject	ID	Nummer	Naam	Polygoon of tracé (GIS) - clic
	Vul eerst tabblad 'gegevens algemeen' in!	1		

Kunnen de projectkosten worden onderverdeeld in subprojecten?

Ja

**! Zorg dat de Macro's in het Excel zijn ingeschakeld (Enable contents) !**

**Toelichting**  
Zie ook "onderverdeling netwerk - Detail niveaus" in tabblad "Scope en uitleg"

#### Netwerkgegevens deel/subprojecten

Optioneel, alleen wanneer kosten data beschikbaar

Deelproject	ID	Nummer	Naam	Polygoon of tracé (GIS) - clic

Gegevens hoofdproject

**#N/A**

Project	#N/A
Project ID	#N/A
Detailnummer	#N/A

**Onderdeel A**  
 project eigenschappen

**A.1-Locatie**

Latitude [deg]	Longitude [deg]

Vergeet niet om een afbeelding van het tracé (GIS / technische tekening o.i.d.) mee te sturen met dit excel!

**A.2-Planning**

	Startdatum	Einddatum
Werkelijk		
Gepland		

**A.3-Grondvervuiling**
*Geef aan op hoeveel meter van het tracé grondvervuiling wel of niet een rol heeft gespeeld.*

Categorie vervuiling	Meter tracé [m]
Onbekend	
	0 m

**Totale tracé lengte:**
**A.4-Niet gesprongen explosieven (NGE)**

Zijn er tijdens de uitvoering additionele en/of voorzorgsmaatregelen getroffen omdat er mogelijk NGE aanwezig waren?

**A.5-Archeologie**

Zijn er tijdens de uitvoering additionele en/of voorzorgsmaatregelen getroffen omdat er archeologische vondsten gedaan zijn?

**A.6-Meekoppelen andere infra**

Zijn er tijdens de aanleg van het warmtenet gelijktijdig andere werkzaamheden uitgevoerd (bijvoorbeeld het vervangen van andere infra)?





### C.1-Project kosten

Wat zijn de totale aanlegkosten van het warmtenet geweest?  
Hoeveel hiervan zijn materiaalkosten?

!	Totale kosten	
	Materiaal kosten	
	Aanlegkosten zonder materiaal	

Zijn deze materiaalkosten inclusief opslag (voor bijvoorbeeld risico en winst)?  
Zijn deze materiaalkosten inclusief levering en transportkosten?

Ja
Ja

*Toelichting: Voor het model zijn we opzoek naar de kosten die een (significante) afhankelijkheid hebben met omgevingsfactoren. Vrij vragen de materiaalkosten (inkoop van leidingen, koppelstukken, nieuwe wegbekeleding, installaties, etc.) apart uit, omdat de aanname is dat deze kosten voornamelijk afhankelijk zijn van de project- en netwerk eigenschappen. Onder materiaalkosten vallen géén materieel kosten die voor de aanleg zijn gebruikt.*

### C.2-Kosten scope

#### Ontwerpkosten

Vul eerst de totale kosten in!

Ja
----

*Toelichting: Onder ontwerpkosten verstaan wij de kosten die gemaakt zijn voor het basisontwerp, definitief ontwerp, en detailontwerp. De ontwerpkosten vallen buiten deze scope, om te garanderen dat de kosten niet worden meegenomen in de analyse vragen wij deze los uit. Als u ze al heeft afgesplitst van de totale projectkosten hoeft u ze hier niet nogmaals op te geven. Kleinschalige ontwerpwerkzaamheden die plaatsvinden tijdens de aanleg van het netwerk hoeven niet los worden opgegeven.*

Voor welke ontwerpwerkzaamheden zitten de kosten inbegrepen?  
Als mogelijk, geef de kosten voor het ontwerp van het net los op.

Ontwerpkosten	
Aanlegkosten zonder ontwerp	

#### Afleversets

Vul eerst de totale projectkosten in!

Ja
----

*Toelichting: De afleversets vallen buiten de scope (scope = aansluitleiding tot achter de voordeurmeterkast begane grond), om te garanderen dat de kosten niet worden meegenomen in de analyse vragen wij deze los uit. Als u ze al heeft afgesplitst van de totale projectkosten hoeft u ze hier niet nogmaals op te -----*

Geef de materiaalkosten van de afleversets los op.

Aanschaf afleversets	
Installatie afleversets	

Als mogelijk, geef de installatiekosten van de afleversets los op.

Aanlegkosten zonder ontwerpkosten en afleversets

#### Stations

Vul eerst de totale projectkosten in!

Ja
----

*Toelichting: Onder station kosten verstaan we de installatiewerkzaamheden, materiaal aanschaf en grondaankoop van de pompstations, warmteoverdrachtstations, regelstations, mengstations, etc. Die aanlegkosten van stations worden wel meegenomen in de analyse, om te garanderen dat de scope van de aangeleverde kosten duidelijk is vragen wij deze los uit.*

Als mogelijk, geef de materiaalkosten van de stations los op.

Materiaal stations	
Installatie stations	
Aankoop grond	

Als mogelijk geef de installatiekosten van de stations los op.

Als er grond is aangekocht, geef dat dan hier op.

Aanlegkosten zonder ontwerpkosten, afleversets en aanleg onderstations

#### Opmerkingen

Zijn er specifieke aandachtspunten m.b.t. dit project waar de invulvelden geen ruimte voor bieden? Benoem ze dan hier.

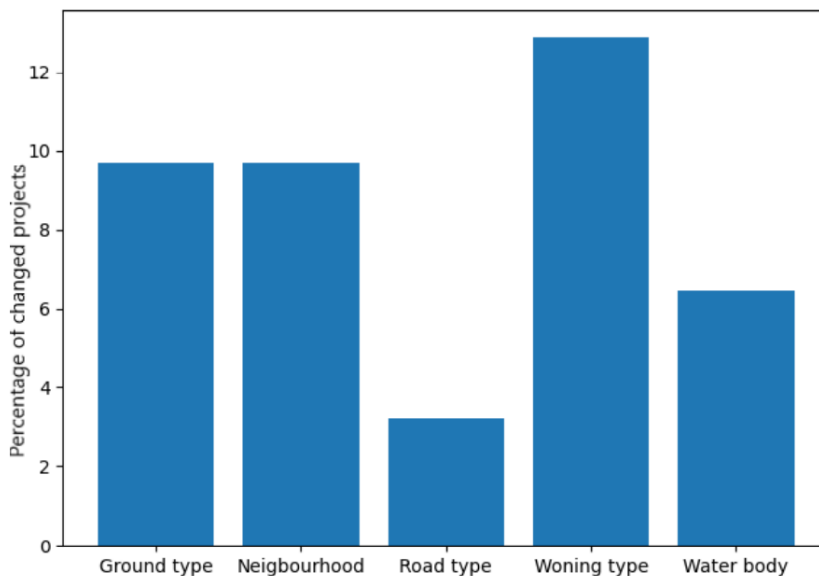
### C. Gevoeligheidsanalyse grootte projectzone

Uit de afstudeer scriptie van Jorn Mieras (Mieras, 2021).

The sensitivity analysis is conducted using 31 projects and 5 different buffer sizes namely 15m, 25m, 30m, 35m, and 45m. After generating all the clipped data layers for the 5 different buffer sizes it became clear that 15m was too small. Because for 12 out of the 31 projects not a single house was inside the buffer area. So in the rest of the sensitivity analysis, only the other 4 buffer sizes are considered.

#### Categorical values

does it matter if the buffer-size is 25m, 30m, 35m, or 45m, or does the category value stay the same no matter which buffer size is chosen. The results of this analyses for 5 considered categorical surrounding parameters can be seen in Figure 4.10. It is important to realize that the vertical axes in the figure represent the percentage of projects that changed and not the number of projects that changed. This means that for most projects (approximately 90 %) the category values of the surrounding parameters do not change when the buffer size is varied between 25m and 45m.



#### Numerical values

For the numerical values, the percentage difference between the minimum and the maximum value out of the 4 different buffer-size is analyzed. As can be seen in the figure for half of the considered surrounding parameters changing the buffer size significantly influences the outcome. For two of these parameters that makes a lot of sense since they are directly dependent on the buffer size. Both building and the road footprint are calculated by dividing respectively the building and road area with the buffer area. The fact that the footprints change that much when the buffer size changes is therefore logical but it does not give any indication on which buffer size is the best. The final surrounding parameter that seems to change significantly when the buffer size is changed is the number of inhabitants. However, when taking a closer look at the projects where the variation is really big the conclusion was drawn that the percentage difference is so big because those projects have a lot more inhabitants than the average number of inhabitants. So when you would look at the percentage difference between the 4 buffer sizes and divide it by the average of

these four different buffer sizes instead of the average of all projects the percentage differences for these projects are a lot smaller. When looking by hand at the differences in the number of inhabitants of different projects the conclusion was drawn that the amount of inhabitants indeed is dependent on the buffer size but not as strong as the boxplot in Figure 4.11 would suggest. The other three surrounding parameters don't vary that much when the buffer size is changed. In the zoomed-in plot in Figure 4.12 it can clearly be seen that the percentages are very low

Based on the sensitivity analysis the conclusion can be drawn that for most surrounding parameters that were checked the size of the buffer does not influence the results significantly. However, since this analysis was only based on a segment of all the projects used in the final model and also not all the surrounding parameters were analyzed, it is decided to do another sensitivity analysis after the final model is developed. In this analysis, all projects are considered but only the surrounding parameters that made it to the final model are checked. The results of this analysis are presented in the discussion. Because of this chosen strategy and because most surrounding parameters are not influenced that much choosing a buffer size by hand by studying some examples projects was deemed a reasonable solution. After studying a couple of example projects it was concluded that 30m was the most logical buffer size. In the example shown in Figure 4.14 it can be seen that the 30m buffer did the best in selecting houses that were facing the trace. Of course, not all examples were as clear as this one when looking at the project in Figure 4.13 for example it can be seen that for this project 25m would have been a better buffer size. However, in this project, 30m is still reasonable where ass 35 is clearly too big. After analyzing 10 example projects it was concluded that 30 meters maybe was not the best solution all the time but it did provide a reasonable buffer zone for all of them and is therefore the best overall buffer size.

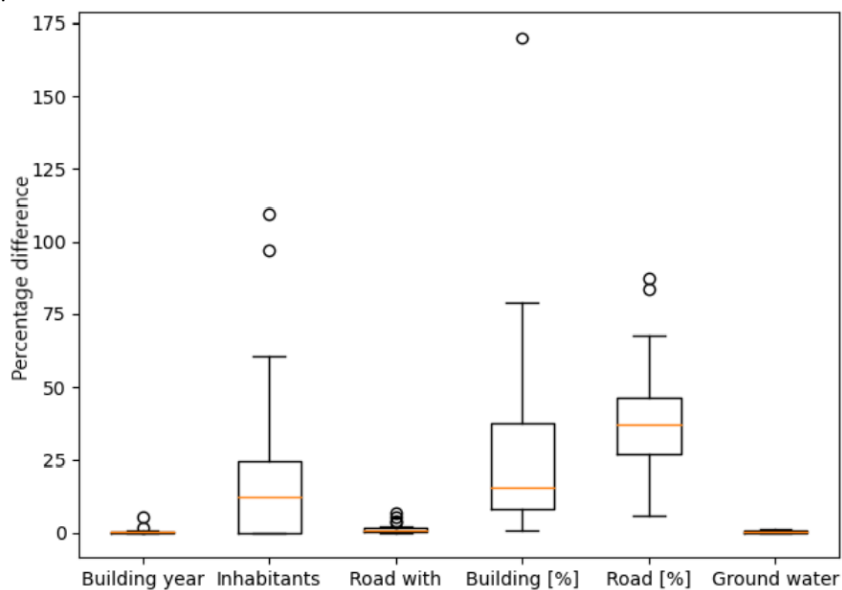




Figure 4.13: Example chosen buildings with different bufferzones



Figure 4.14: Example 2 chosen buildings with different bufferzones

## D. Model resultaten - Uitvoeringskosten

### Model Building Summary

Target: Uitvoeringskosten

	Model									
	1	2	3	4	5	6	7	8	9	10
Adjusted R Square	,862	,860	,859	,858	,858	,858	,858	,858	,858	,857
Woongebied	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bemaling_transformed	✓		✓		✓	✓		✓	✓	
Postcode_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Type_transformed	✓	✓	✓	✓	✓	✓	✓		✓	✓
Grondvervuiling_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lengte_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Voetafdruk_Weg_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Effect Tijdstempel_transformed	✓	✓	✓		✓		✓	✓	✓	✓
Supply_temp_max_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Aantal_afnemers_transformed	✓	✓		✓	✓	✓		✓		✓
Aantal_aansluitingen_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Leiding_diameter_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lengtevervuldtrace_transformed	✓	✓	✓	✓		✓	✓	✓		
Peilgebied_niveau_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bouwjaar_transformed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

The model building method is Best Subsets using the Adjusted R Square criterion.  
A checkmark means the effect is in the model.

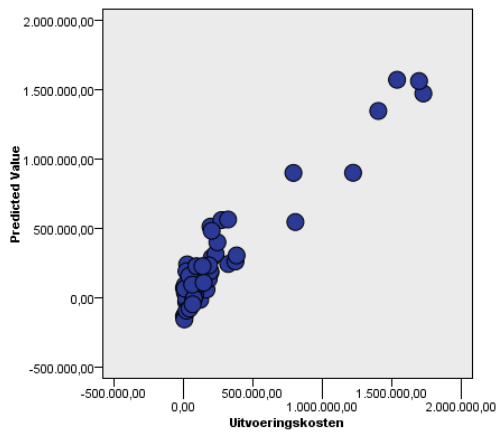
## Automatic Data Preparation

Target: Uitvoeringskosten

Field	Role	Actions Taken
(Aandeelvervuildegrondvantaal_transformed)	Predictor	Trim outliers Replace missing values
(Aangelegd_vermogen_transformed)	Predictor	Trim outliers Replace missing values
(Aangesloten_vermogen_transformed)	Predictor	Trim outliers Replace missing values
(Aantal_aansluitingen_transformed)	Predictor	Trim outliers Replace missing values
(Aantal_afnemers_transformed)	Predictor	Trim outliers Replace missing values
(Afstandussengebouwenweg_transformed)	Predictor	Trim outliers Replace missing values
(Bemaling_transformed)	Predictor	Replace missing values
(Bouwjaar_transformed)	Predictor	Change measurement level from ordinal to continuous Trim outliers Replace missing values
(Delta_temp_max_transformed)	Predictor	Trim outliers Replace missing values
(Diepte_opensleuf_transformed)	Predictor	Trim outliers Replace missing values
(Dominant_leidingtype_transformed)	Predictor	Merge categories to maximize association with target
(Geplandeduur_transformed)	Predictor	Trim outliers Replace missing values
(Grondvervuiling_transformed)	Predictor	Replace missing values Merge categories to maximize association with target
(Grondwater_niveau_transformed)	Predictor	Trim outliers
(Inwonersdichtheid_transformed)	Predictor	Trim outliers
(Leiding_diameter_transformed)	Predictor	Trim outliers Replace missing values
(Lengte_transformed)	Predictor	Trim outliers
(Lengtevervuiltrace_transformed)	Predictor	Trim outliers Replace missing values
(Peilgebied_niveau_transformed)	Predictor	Trim outliers Replace missing values
(Postcode_transformed)	Predictor	Merge categories to maximize association with target
(Standaarddeviatie_Bouwjaar_transformed)	Predictor	Trim outliers Replace missing values
(Supply_temp_max_transformed)	Predictor	Trim outliers Replace missing values
(Tijdstempel_transformed)	Predictor	Trim outliers Replace missing values
(Trefkans_archeologie_transformed)	Predictor	Trim outliers Replace missing values
(Type_transformed)	Predictor	Replace missing values Merge categories to maximize association with target
(Voetafdruk_Bomen_transformed)	Predictor	Trim outliers
(Voetafdruk_gebouw_transformed)	Predictor	Trim outliers Replace missing values
(Voetafdruk_Weg_transformed)	Predictor	Trim outliers
(Voetafdruk_Wegtype_kruising_transformed)	Predictor	Trim outliers Replace missing values
(Voetafdruk_Wegtype_overigverkeersgebied_transformed)	Predictor	Trim outliers Replace missing values
(Voetafdruk_Wegtype_verbinding_transformed)	Predictor	Trim outliers Replace missing values
(Wegbreedte_transformed)	Predictor	Trim outliers
Wegbreedte_2tot4m	Predictor	Exclude predictor: constant

If the original field name is X, then the transformed field is displayed as (X\_transformed). The original field is excluded from the analysis and the transformed field is included instead.

**Predicted by Observed**  
Target: Uitvoeringskosten



**Outliers**  
Target: Uitvoeringskosten

Record ID	Uitvoeringskosten	Cook's Distance
3	195.307,00	0,847
45	1.728.500,00	0,308
59	1.222.360,00	0,266
62	806.686,00	0,165
60	203.508,00	0,109
49	274.220,00	0,098
33	231.916,00	0,095

Records with large Cook's distance values are highly influential in the model computations. Such records may distort the model accuracy.



## E. Model resultaten - M\_Uitvoeringskosten

### Model Building Summary Target: M\_uitvoeringskosten

	Model									
	1	2	3	4	5	6	7	8	9	10
<b>Adjusted R Square</b>	,559	,558	,556	,555	,555	,554	,547	,547	,544	,544
<b>Waterlichaam</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Postcode_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Wegbreedte_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Voetafdruk_Weg_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Inwonersdichtheid_transformed</b>	✓	✓	✓			✓			✓	
<b>Voetafdruk_Bomen_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Effect Supply_temp_max_transformed</b>	✓		✓	✓	✓				✓	✓
<b>Delta_temp_max_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓		
<b>Diepte_opensleuf_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Aandeelvervuildegrondvantotaal_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Voetafdruk_Wegtype_kruising_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<b>Trefkans_archeologie_transformed</b>	✓	✓			✓			✓	✓	
<b>Bouwjaar_transformed</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

The model building method is Best Subsets using the Adjusted R Square criterion.  
A checkmark means the effect is in the model.

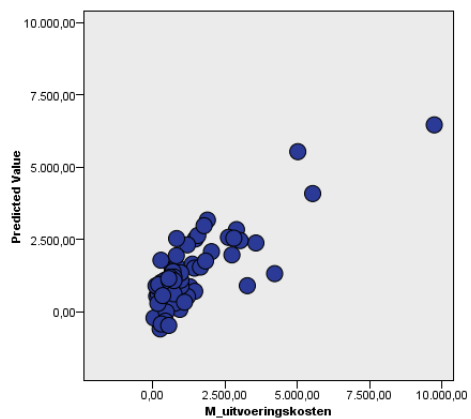
## Automatic Data Preparation

Target: M\_uitvoeringskosten

Field	Role	Actions Taken
(Aandeelevuildegrondvantotaal_transformed)	Predictor	Trim outliers Replace missing values
(Aangesloten_vermogen_transformed)	Predictor	Trim outliers Replace missing values
(Aantal_aansluitingen_transformed)	Predictor	Trim outliers Replace missing values
(Aantal_afnemers_transformed)	Predictor	Trim outliers Replace missing values
(Afstandussengebouwenweg_transformed)	Predictor	Trim outliers Replace missing values
(Bemaling_transformed)	Predictor	Replace missing values
(Bouwjaar_transformed)	Predictor	Change measurement level from ordinal to continuous Trim outliers Replace missing values
(Delta_temp_max_transformed)	Predictor	Trim outliers Replace missing values
(Diepte_opensleuf_transformed)	Predictor	Trim outliers Replace missing values
(Geplandeduur_transformed)	Predictor	Trim outliers Replace missing values
(Grondverbetering_transformed)	Predictor	Replace missing values
(Grondwater_niveau_transformed)	Predictor	Trim outliers
(Inwonersdichtheid_transformed)	Predictor	Trim outliers
(Leiding_diameter_transformed)	Predictor	Trim outliers Replace missing values
(Lengte_transformed)	Predictor	Trim outliers
(Lengtevervuldtrace_transformed)	Predictor	Trim outliers Replace missing values
(Peilgebied_niveau_transformed)	Predictor	Trim outliers Replace missing values
(Postcode_transformed)	Predictor	Merge categories to maximize association with target
(Standaarddeviatie_Bouwjaar_transformed)	Predictor	Trim outliers Replace missing values
(Supply_temp_max_transformed)	Predictor	Trim outliers Replace missing values
(Tijdstempel_transformed)	Predictor	Trim outliers Replace missing values
(Trefkans_archeologie_transformed)	Predictor	Trim outliers Replace missing values
(Voetafdruk_Bomen_transformed)	Predictor	Trim outliers
(Voetafdruk_gebouw_transformed)	Predictor	Trim outliers Replace missing values
(Voetafdruk_Weg_transformed)	Predictor	Trim outliers
(Voetafdruk_Wegtype_kruising_transformed)	Predictor	Trim outliers Replace missing values
(Voetafdruk_Wegtype_overigverkeersgebied_transformed)	Predictor	Trim outliers Replace missing values
(Voetafdruk_Wegtype_verbinding_transformed)	Predictor	Trim outliers Replace missing values
(Wegbreedte_transformed)	Predictor	Trim outliers
Wegbreedte_2tot4m	Predictor	Exclude predictor: constant

If the original field name is X, then the transformed field is displayed as (X\_transformed). The original field is excluded from the analysis and the transformed field is included instead.

**Predicted by Observed**  
Target: M\_uitvoeringskosten



**Outliers**  
Target: M\_uitvoeringskosten

Record ID	M_uitvoeringskosten	Cook's Distance
11	9.729,21	1,351
51	821,26	0,297
30	5.522,19	0,105
8	1.470,26	0,092
34	4.211,05	0,092
62	3.265,94	0,084

Records with large Cook's distance values are highly influential in the model computations. Such records may distort the model accuracy.