Final report

Optimisation of the Production Distribution over the Groningen field to reduce Seismicity

1 December 2017

Table of Contents

Management Summary

This document describes the potential for optimisation of the distribution of the gas production over the clusters of the Groningen field, aiming to reduce seismicity and risk. The optimisation results are also presented and analysed. Based on this analysis of the optimisation results, the impact of adjustments to the current distribution of the production over the clusters in the field will be reviewed. Assumption in all the analyses is the currently allowed production of the Groningen field, 21.6 N Bcm per gas year.

Following the Instemmingsbesluit Winningsplan 2013, five production clusters in the Loppersum area were closed-in as a first step in the optimisation of the production distribution to reduce seismicity. In Winningsplan 2016 NAM presented an alternative distribution of the production, aiming to further reduce the seismicity based on insights gained in the assessment of Hazard and Risk.

The optimisation tool and methodology as used in this optimisation make use of the Groningen reservoir model linked to the hazard and risk assessment model for induced seismicity in Groningen. Two methods have been used; (1) an extension of the Groningen reservoir model to also include compaction and seismicity based on a proxy approach and (2) a direct coupling of the Groningen reservoir model to the hazard and risk model. This kind of mathematical optimisation has never been tried before, is at the forefront of what is technically possible and requires access to large computer power.

The optimisation of the distribution of the production over the field to reduce seismicity, can be used to minimise different metrics of seismicity, depending on the preferred optimisation goal. For this analysis three goals for the next 5-year period (2018-2022) were used:

1. Reduction of nuisance: To achieve a minimisation of nuisance, the number of earthquakes can be minimised.
2. Reduction of the hazard: To reduce hazard, the largest PGV or PGA can be minimised.
3. Reducing of risk for the total population: To reduce the risk for the total population, the population weighted PGV can be minimised.

In this document, different optimisations, aiming to minimise these different seismicity metrics are presented.

Apart from minimizing the seismicity metrics, the optimisation will also (re-)distribute the epicentre locations of the earthquakes differently over the Groningen field area. In other words: with every upside there is also a some downside in seismic effects for a certain part of the Groningen field. When production is reduced in one area of the Groningen field and the production as a whole stays at the current level, this inherently will lead to an increase in production in other parts of the field. The optimisation will therefore impact differently on individuals and local communities in the area.

Each optimisation goal will result in a different optimised distribution of the production over the clusters, which will cause the optimisation to impact differently on individuals and local communities in the area. For instance, if the optimisation is carried out to minimise risk for the total population living above the field, using population weighted PGV, people living in urban environments in close proximity to their neighbours will individually benefit more than people living in rural areas living further apart.

An important consequence of these findings is that NAM cannot advise as to what is the best optimisation of the production. In this document, NAM will outline the scientific findings of the research that has been done, outline the possible seismicity metrics and give insight in the consequences of each possible optimisation goal. It is up to the authorities to select the preferred optimisation goal and what result is best for the Groningen field and population as a whole. NAM does emphasise that in all presented options, the seismic risk is within the Meijdam norm as set by the Minister.

Results

Analysis of the optimisation results shows that reducing production from some selected clusters, while redistributing that production volume over the other clusters, can reduce the number of earthquakes or risk for the population in that area. As mentioned before, this does come with a downside for part of local community, being the part of the population that is living in the vicinity of the clusters where the production is increased.

From the different optimisation trials, the following conclusions can be drawn:

- Reduction of the production from the Eemskanaal cluster and redistribution of the production from these clusters over the other clusters, is expected to result in a lower number of earthquakes and reduce risk for the total population. Both optimisations aiming to reduce the number of earthquakes and those aiming to reduce risk for the total population, show this same result.
- Increase of production from the ‘t Zandt cluster, allowing production from other clusters to be reduced, is expected to result in a decrease of risk for the total population. The impact of variation of the production from ‘t Zandt cluster is neutral for the number of earthquakes.
- Increase of production from the Leermens cluster, allowing production from other clusters to be reduced, has a neutral impact on risk for the total population and increases slightly the total number of earthquakes.
- Minimisation of production from the Ten Post, De Paauwen and Overschild clusters benefits both the reduction in the number of earthquakes and the total population risk.

In order to quantify the results of the optimisations, NAM has made analyses that show what percentage of the different goals can be achieved over the next 5-year period (2018 – 2022) compared to the current situation:

- When optimizing for the number of earthquakes, the model suggests a reduction by 15% can be achieved. The associated impact on population weighted PGV for this optimisation is 12%.
- When optimizing for population weighted PGV, the model suggests a reduction of 15% can be achieved. The association impact on the number of earthquakes is 10% for this optimisation.
- A hybrid optimisation, aiming to achieve a combined minimisation of both the risk for the population and number of earthquakes, was also investigated.

In conclusion, a further optimisation in order to meet any of the three goals is possible by closing in the Eemskanaal cluster and balancing this production by increased production from clusters located in the north and east of the field. Comparing the production distribution resulting from the three different optimisations shows that different clusters are needed to balance production. If the aim is to reduce the

risk for the total population (goal 3) production from the 't Zandt cluster and possibly the Leermens cluster should be increased. If the aim is to reduce the number of earthquakes (goal 1), only the 't Zandt cluster should be opened up. Production from the clusters Ten Post, De Paauwen and Overschild, should remain minimized.

The optimisation results are in overall agreement with the optimisation carried out in the first quarter of 2016 and reported in the technical documentation for Winningsplan 2016 (Ref. [1]). This earlier optimisation was primarily based on the practical understanding of the response of the pressure in the reservoir and seismicity over the field to production changes. That both methods for optimisation give very similar results, lends confidence to these results.

Implementing Optimisation on the basis of the results

When deciding as to what optimisation should be implemented on the basis of these model-based optimisation results, the authorities are advised to take into account some practical and operational considerations and limitations:

- Impact on capacity from the field (security of supply, interest relevant to the Minister) and ability to achieve flat production (advice from SodM) from the different areas of the field.
- Ability to implement the production distribution in light of operating constraints like the custody transfer stations at which the gas needs to be delivered to GTS, and the requirement to fill up the Norg gas storage facility.
- A Meet- en Regelprotocol is in effect to ensure monitoring of the field and operational response to the changes in seismicity or unexpected events. There could be interference between the current Meet en Regelprotocol and a new regional production distribution – e.g. triggering local seismicity thresholds. It should be considered in advance how these interferences are handled.

It is recommended that any adaptations to the distribution of the production from the field should be implemented gradually, while closely monitoring the impact on seismicity. Within the framework of the Meet- en Regelprotocol, the effectiveness of the production measures will be regularly reviewed, when more data will have become available. In response, the optimisation will need to be updated regularly. A dedicated test plan to gather data on the response of the seismicity to production changes will be prepared. An example of a potential action in a test plan is the 2017 installation of the geophone string in the Harkstede well to investigate the high seismicity in the Ten Boer area (Ref. [2]), which recorded low seismicity levels in the area. Such an observation could potentially lead to a period of controlled production from this cluster to confirm the model and the effectiveness of the production measure.

These updates will also allow improvements in the optimisation methodology to be implemented. The current model has been calibrated to maximize the model quality and prevent it from being over-determined. NAM prefers to identify and take optimisation measures that might have a positive effect (but do no harm), even if the effect of the measures is currently not yet proven and quantified. Further enhancements of the modelling- and optimisation methods and calibrations to more data are expected to improve the recommended distribution of the production in the future.

Conclusion

State of the art methods are used in providing the results of this report. That does however, and unfortunately, not mean that a simple solution can be presented. The optimised production from the Groningen field involves a delicate balance of sometimes conflicting interests and also of operational restrictions.

NAM has presented the options of the optimisation of the production based on scientific research and has included the impact each change has for the population in Groningen as a whole and also regional differences. We have shown that seismicity in the 2018-2022 can be reduced by redistributing production over de Groningen field. Dependent on the metric to be optimised (nuisance, hazard, risk), different optimised distributions ensue. This represents a choice that needs to be made by policy makers. Taking account of operational limitations, risk reduction or nuisance reduction of more than 10% is possible. This risk reduction is relatively efficient, e.g. in comparison with total field volume reductions that would achieve similar reductions in seismicity.

We are looking forward to the Minister's response as to the weighing of these interests and the resulting conclusions concerning the optimisation of the production of the Groningen field.

Based on a review of the spatial extent of pressure transients at different timescales and for recent production history, we conclude that it is not feasible to optimise production fluctuations at individual production clusters over short timespans (hours-weeks). Over longer timespans (months) and regions (groups of multiple production clusters) production fluctuations can be avoided. In the operationalization section of this report (Chapter 7) the regions have been chosen to achieve this.

1 Introduction

1.1 Instemmingsbesluit

As part of the 30/9/2016 Instemmingsbesluit WP2016 (and of the subsequent 24/5/2017 Wijzigingsbesluit), NAM was tasked in Article 3.2 to investigate whether an alternative distribution of production from the Groningen field could reduce the seismic hazard or risk. The text of the Wijzigingsbesluit reads:

> **Artikel 3**
>
> 2. De Nederlandse Aardolie Maatschappij B.V. onderzoekt of een alternatieve verdeling van de productie over alle regio's tot een lagere seismische dreiging of seismisch risico leidt en brengt daarover uiterlijk op 1 december 2017 ten genoegen van de inspecteur-generaal der mijnen, een rapport uit aan de Minister van Economische Zaken. Daarbij wordt tevens in detail voor alle clusters onderzocht of clusteroperaties verder geoptimaliseerd kunnen worden ten einde (ook regionale) fluctuaties zo veel mogelijk te beperken. De Nederlandse Aardolie Maatschappij B.V brengt uiterlijk op 1 februari 2017 een plan van aanpak in bij de inspecteur-generaal der mijnen. Uiterlijk op 1 oktober 2017 dient de Nederlandse Aardolie Maatschappij B.V. ter beoordeling van de inspecteur-generaal der mijnen een concept-rapportage in. Een eventuele alternatieve productieverdeling wordt niet ingevoerd voordat deze ten genoegen van de inspecteur-generaal der mijnen is.

The requested Plan of Approach (Plan van Aanpak) and draft report (concept-rapportage) have been shared with SodM prior to the due dates in the Instemmingsbesluit. The current document is the final report with a due date of 1st December 2017.

Article 5 of the Instemmingsbesluit addressed the Meet- en Regelprotocol (Measurement and Control protocol). In article 5e, a methodology for the optimisation of the distribution of production aiming to reduce seismic risk was requested.

> **Artikel 5**
>
> 1. De Nederlandse Aardolie Maatschappij B.V. dient uiterlijk op 1 juni 2017 bij de Minister van Economische Zaken een nieuw Meet- en regelprotocol in, waarin tot genoegen van de inspecteur-generaal der mijnen wordt beschreven:
>
> e. de methodiek waarmee een optimale verdeling van de productie uit oogpunt van seismisch risico wordt bepaald.

The report describing the methodology was submitted to SodM on the 29th May 2017, Reference [3].

1.2 Previous Optimisation Studies

In the first quarter of 2016, the first optimisation of the distribution of the production from the clusters in the Groningen field to reduce seismicity was carried out. This was reported in the Technical Addendum to the Winningsplan submitted on 1st April 2016, Reference [1]. This initial optimisation was based both on mathematical optimisation and on a practical understanding of the pressure response of the reservoir

to the distribution of the production and the impact on the resulting seismicity. The impact of the alternative distribution of production on the seismicity and risk was checked by performing dedicated hazard and risk assessments.

The current optimisation of the distribution of the production with the aim to reduce seismicity, was carried out with more mathematical rigour and based on the latest study insights and all available data.

1.3 Groningen Production System

There are 22 production locations (20 clusters and two satellites) by which gas can be produced from the field (Figure 1-1). The wells of an individual cluster typically produce gas from a small area of the field underneath the cluster area. The only exception is the deviated well, EKL-13, that is producing gas from the reservoir at some distance from the other (vertical) wells of the Eemskanaal cluster. All these clusters can be controlled relatively independently within their respective operational constraints. This offers the capability to make choices in how production is spatially distributed in the field.

Figure 1-1 shows how the historical earthquakes display some degree of spatial clustering. The various regulatory production measures that were imposed to date, have already achieved an optimisation of production distribution.

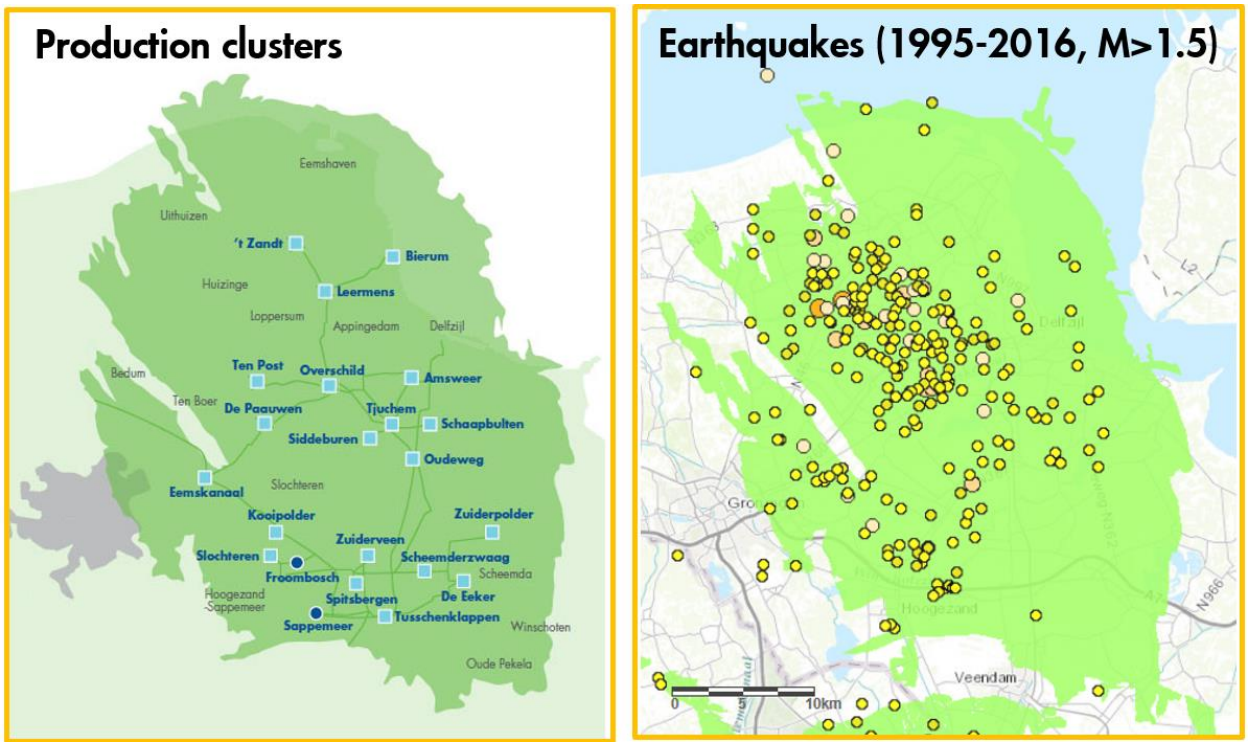


Figure 1-1: Hazard (earthquakes) and controls (production clusters)

1.4 Constraints on Production

The areal distribution of production from the field is to be optimised, aiming to minimise seismicity (Instemmingsbesluit - Article 3.2). Currently, the distribution of the production from the field is limited both by constraints imposed in the Instemmingsbesluit, and by constraints in the production system.

1.4.1 Current Regulatory Constraints

The "Instemmingsbesluit – Winningsplan 2016" and the following Wijzigingsbesluit contain several articles (articles 2, 3 and 4) impacting the production of gas from the Groningen Field (following the numbering of articles in the Instemmingsbesluit):

Article 2.1	Initially, set a Field Cap of 24 N Bcm/gas-year. This was adjusted in article 1 of the Wijzigingsbesluit to 21.6 N Bcm/gas-year.
Article 2.2	The additional volume depending on degree-days up was initially set at 6 Bcm/gas-year (with appendix on calculation of degree –days). In article 2.c of the Wijzigingsbesluit this was adjusted to 5.4 N Bcm/gas-year,
Article 2.3	Additional volume depending on technical issues (transport restriction, failure GTS system and hi-cal composition) up to 1.5 Bcm/gas-year,
Article 2.4	Administration of additional volume failure GTS system,
Article 2.5	Administration of additional volume depending on degree-days
Article 3.1	Regional off-take from the field pro-rated to the regional caps,
Article 3.3	Five Clusters around Loppersum reduced to minimally required volumes for Security of Supply,
Article 4.1	Reduction of seasonal variations and monthly variations; temporal flat production,
Article 4.2	Introduction of production changes.

The regional caps in Article 3.1 refer back to the areal offtake distribution restrictions as imposed on 30/1/2015. Dedicated caps were imposed for the following regions[1]:

LOPPZ[2]:	Stand-by rates for security of supply only (to a maximum of 3 Bcm[3])
Eemskanaal:	2.0 Bcm per year,
South-West:	9.9 Bcm per year,
East:	24.5 Bcm per year.

Figure 1-2 highlights these regulatory production regions on a map of the field, showing how the production clusters are assigned to the various regions.

Note that although the sum of the regional caps exceeds the total field production cap of 24 N Bcm/year in the Instemmingsbesluit, the requirement to maintain a pro-rated regional offtake limits operational flexibility.

[1] All caps are in 100% Wellhead N.m3

[2] The LOPPZ clusters are located in the earthquake prone Loppersum area, and constitute of **L**eermens, **O**verschild, De **P**aauwen, Ten **P**ost and 't **Z**andt

[3] NAM tries to minimize this volume, in 2016 1.0 Bcm was used.

In this study, the constraint imposed by Article 1 of the Wijzigingsbesluit, limiting the annual production of the field to 21.6 N Bcm/year, will be imposed on the production for all optimisation forecasts. The other constraints of the Instemmingsbesluit and Wijzigingsbesluit will not be included in the optimisation as hard limits. However, as these constraints have been imposed to achieve an optimised production distribution, the current optimisation is in effect a study to assess whether these current constraints can be improved upon to achieve a further reduction in the impact of seismicity.

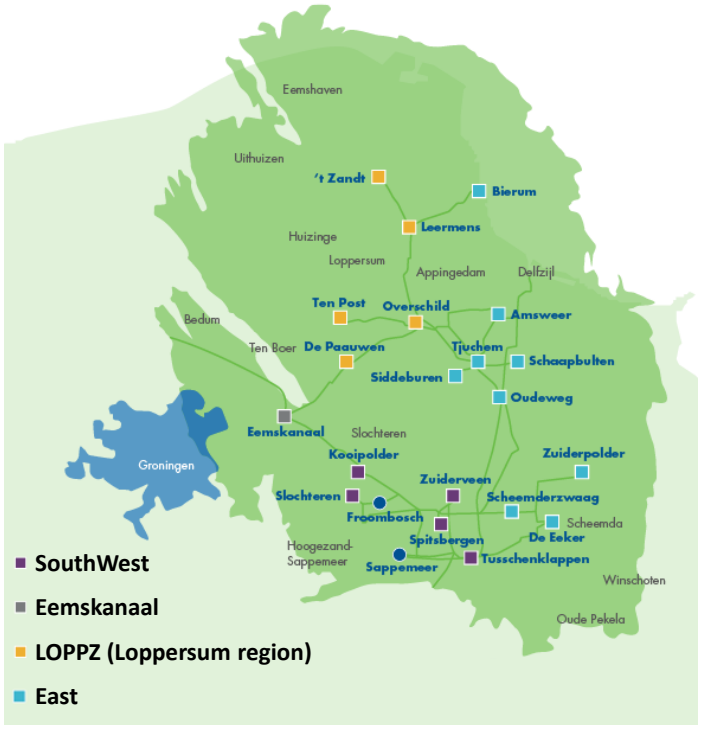


*Figure 1-2: Regulatory production regions: **South-West** (clusters Kooipolder, Slochteren, Zuiderveen, Spitsbergen, Tusschenklappen, Froombosch, Sappemeer), **Eemskanaal** (cluster Eemskanaal), **East** (clusters Bierum, Amsweer, Schaapbulten, Tjuchem, Siddeburen, Oudeweg, Zuiderpolder, Scheemderzwaag, De Eeker), and **Loppersum** (clusters 't Zandt, Overschild, De Paauwen, Ten Post, Leermens)*

1.4.2 Operational Constraints

In addition to the regulatory constraints, there are also technical limitations on production from the Groningen production system. On a high level, the following components make up the Groningen production system (Figure 1-3):

Clusters — There are currently 22 production locations, 20 clusters and two satellites. The 20 production clusters are equipped with compression and gas process facilities to bring the gas to sales gas specification. The satellites are connected to a production cluster (Froombosch is connected to Slochteren, Sappermeer is connected to Tusschenklappen) and not equipped with a compressor and process equipment.

Wells	From some 250 wells, gas is produced in the field. Well capacity is limited and depends on the (declining) reservoir pressure.
Compressors	The compression capacity is limited by the power of the compression drivers and the compressor operating envelope.
Gas process	Produced gas is brought to the gas quality as stipulated in the Gaslaw (Regeling van de Minister van Economische Zaken van 11 juli 2014, nr. WJZ/13196684, tot vaststelling van regels voor de gaskwaliteit (Regeling gaskwaliteit) making use of Joule-Thomson effect.
Ring System	To evacuate the produced gas, clusters are connected to NAM's gas pipeline grid commonly referred to as the 'Groningen ring' but consists of a more complicated configuration consisting of 136 valves and 59 different sections of pipeline (total 162 km). Via the pipeline system gas is distributed from the production clusters over the custody transfer stations.
Custody Transfer Station	By means of 7 custody transfer stations, where gas quality and quantity is measured, the Groningen ring is connected to the GasUnie pipeline grid. Every transfer station feeds one GTS pipeline and could be considered as the starting point of a pipeline. Offtake per custody transfer station is managed by GTS by manipulating pressure in the gastransport pipelines.
Underground Gas Storage	Working-volume produced from the Norg UGS in the production season is reinjected in the injection season. The UGS Norg is connected to the GTS network as well as to the Groningen ring with a dedicated pipeline (NorGron pipeline). This pipeline ties-in to the south-western part of the ring. Gas can be evacuated via the GTS network connection and the NorGron pipeline. Gas is injected via the NorGron pipeline or GTS network connection. The higher the suction pressure of the injection compressors, the more efficient the injection process. The suction pressure can be controlled by injecting via NorGron and by increasing the pressure in the Groningen ring. An increased pressure in the Groningen ring requires a specific operational set-up (also known as ring segregation). Ring segregation is achieved by opening and closing valves in the Groningen ring, by which selection is made in production clusters that feed UGS Norg.

The fraction of the production capacity of this total production system that is available at any point in time is governed by the availability of the system. Scheduled periods for maintenance and unscheduled stops caused by failures, both impact availability.

The Groningen production system cannot be seen in isolation, as it is pivotally linked to the functioning of the gas supply system in the Netherlands (and parts of Belgium, France and Germany). Operational changes (can) have immediate impact on the transport system operation.

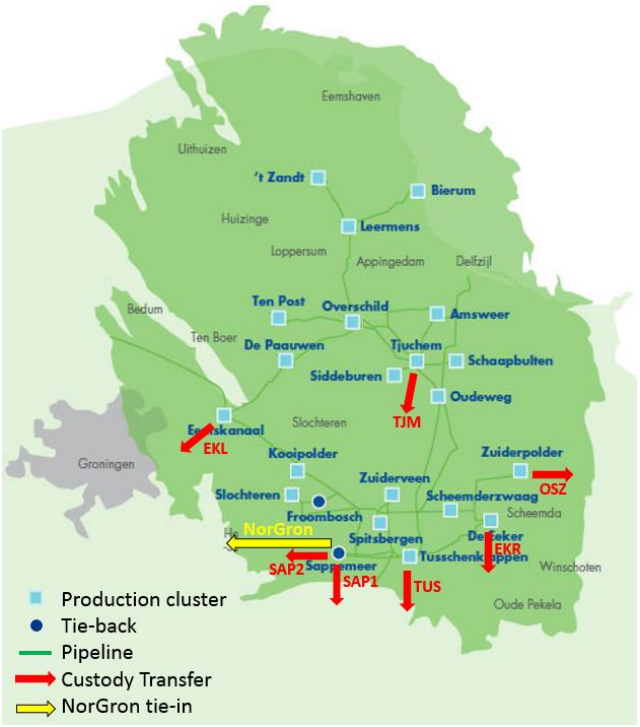


Figure 1-3 Groningen production system

1.5 Optimisation scope

A first control measure to optimise the distribution of production with the aim to reduce seismicity was imposed by the Minister of Economic Affairs. On 17/1/2014 the total offtake from the five production clusters within the seismicity prone Loppersum area was reduced to a maximum of 3 Bcm per annum. This measure was successful in reducing seismicity in this area in the years following the implementation of the measure, see Reference [4]. Any optimisation will in effect be an increment to this already successful optimisation of the distribution of production over the field.

1.5.1 Optimisation of Distribution Production in Winningsplan 2016

As part of the Winningsplan 2016 submission, NAM investigated an alternative expert judgement optimisation of the offtake across the various production clusters in order to minimize the seismic risk, Reference [1]. As part of this work an observation was made with respect to the field offtake as dictated by the regulatory constraints as of November 2015. The predicted risk associated with this offtake showed that the area of highest risk roughly coincides with the band Bedum – Loppersum – Appingendam – Delfzijl, and that in the North-East of the field there is an area which is poorly drained, but imposes more limited risk (Reference [1] and Figure 1-4). A production distribution scenario was formulated to increase offtake from this area (North-East of the field), and decrease the offtake from the Eemskanaal cluster (located towards Groningen city), while keeping the same offtake from the full field.

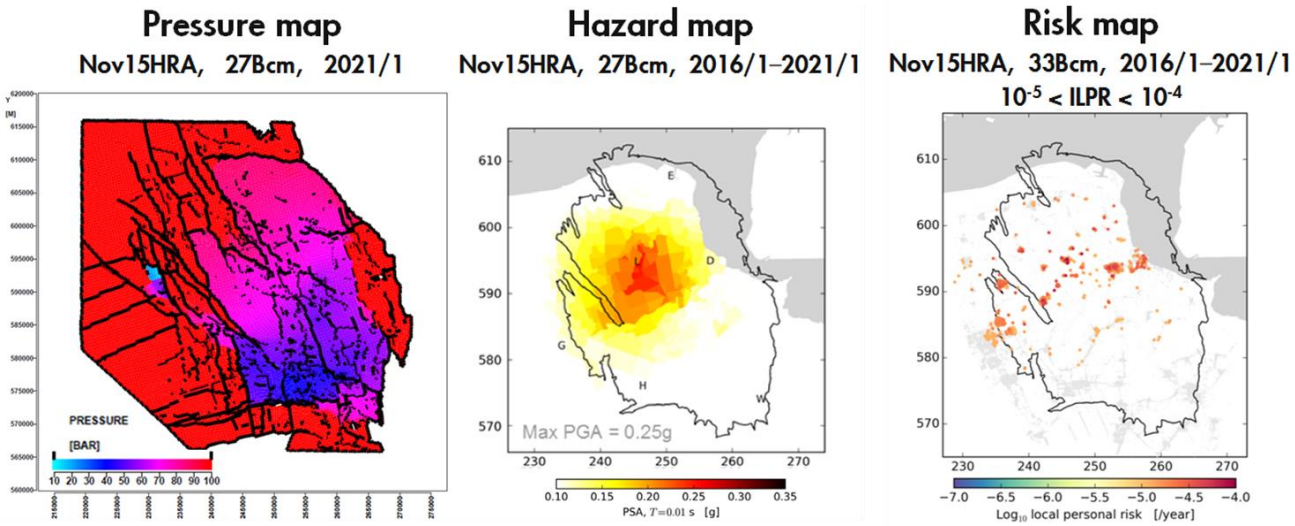


Figure 1-4 Combined visualization showing 1) the expected reservoir pressure distribution across the field in 2021 as per the Nov2015HRA forecast at 27Bcm/y ; 2) the associated Hazard map ; 3) the Risk map at 33Bcm annual offtake

This distribution was not chosen as the base case production scenario in the Winningsplan 2016, because at the time the analysis was less comprehensive and the effect of the incremental optimisation on seismicity was assessed to be relatively small.

1.5.2 Impact of production redistribution on reservoir pressure

The models, as used in this analysis, relate pressure depletion to compaction, and compaction to seismicity. Hence redistribution of pressure depletion rate (for a given time window) will alter the modelled seismic response of the system.

The Groningen reservoir consists of high quality reservoir rock with high net-to-gross, thick column, and good permeability. The field is heavily faulted with more than 1,100 major and minor faults identified on seismic. Across most faults there is a good sand-to-sand juxtaposition. Therefore, many of the faults do not act as major baffles to gas flow. Consequently, there is a mostly good pressure communication across the field. However, due to the size of the field, reducing the offtake in one region will cause a pressure imbalance. Initially, the pressure decline in a low offtake area will slow down. But over time, when the pressure imbalance at the field scale becomes larger, this becomes a driving force causing gas in the higher pressured region to flow towards the lower pressure regions. This effect will be seen over distances of many kilometres, because of the good pressure communication across the field. The pace at which the pressure equilibration process takes place is slow because of the high compressibility of gas at reservoir conditions. Pressure equilibration can take several years. Furthermore, a few large faults with limited gas-to-gas juxtaposition will regionally act as baffles. Examples are the NW-SE trending major faults to the north-west of the Loppersum area.

From Figure 1-5 it can be observed that the regulatory imposed production measures in the Loppersum region have been successful to (temporarily) arrest the pressure decline (and hence compaction) in that region. The imposed production cap in January 2014 reduced the production rate at Loppersum by some 90%, which temporarily reduced the pressure decline in that area almost completely. However, as a result

of the imbalance in offtake rates across the field, a pressure gradient formed from the North down to the South of the field, as evident from Figure 1-5 and Figure 1-6. Because the entire field is hydraulically connected, this gradient forms the driving force for gas to start flowing from the North towards the South, which eventually causes the pressure decline in the Loppersum area to fall back to the field average pressure decline once a pseudo-steady state pressure field is established across the field for constant production rates. Figure 1-8 shows the current drainage patterns using streamlines[4]. It can be observed that the Loppersum area is currently being drained by several other cluster regions, including the most southern ones as well as those to the east.

> Note that in these figures, various areas of the field are annotated. These areas will be placed in context later on in the report (chapter 5), but are already defined in Figure 1-7 for reference. The cluster abbreviations are defined in Appendix A.

The various cluster production rates across the field govern the magnitude of the pressure gradient. More production from the South and/or less production from the North will trigger another transient response, which will start to increase the gradient until eventually a new pseudo-steady state forms (should such new production rates remain constant). Any transient response will however be proportional to the magnitude of the change in production rates. Given that the production rates in the Loppersum area were already reduced by 90%, reduction of the remaining 10% down to zero will only have a limited impact. On the other hand, the production clusters in the South of the field are already running at high load factors, with limited scope for further increase. At some point the production compressors will run at their limits, and rates will drop with ongoing reservoir pressure depletion. In the current production regime, this will happen first in the South of the field.

[4] These streamlines were calculated in the dynamic reservoir simulator as a post-processing step, based on the pressure field of the active timestep. They basically visualize the pressure gradients, which govern/reflect the flow path of gas molecules. It is important to realize that the streamlines as visualized in Figure 1-8 do not address the travel times of individual gas molecules, and as such should not be misunderstood: they do not reflect an instantaneous drainage. As described in chapter 9, due to the dampening effect of a highly compressible fluid flow in a porous medium, the pressure gradients further away from the production clusters are tiny.

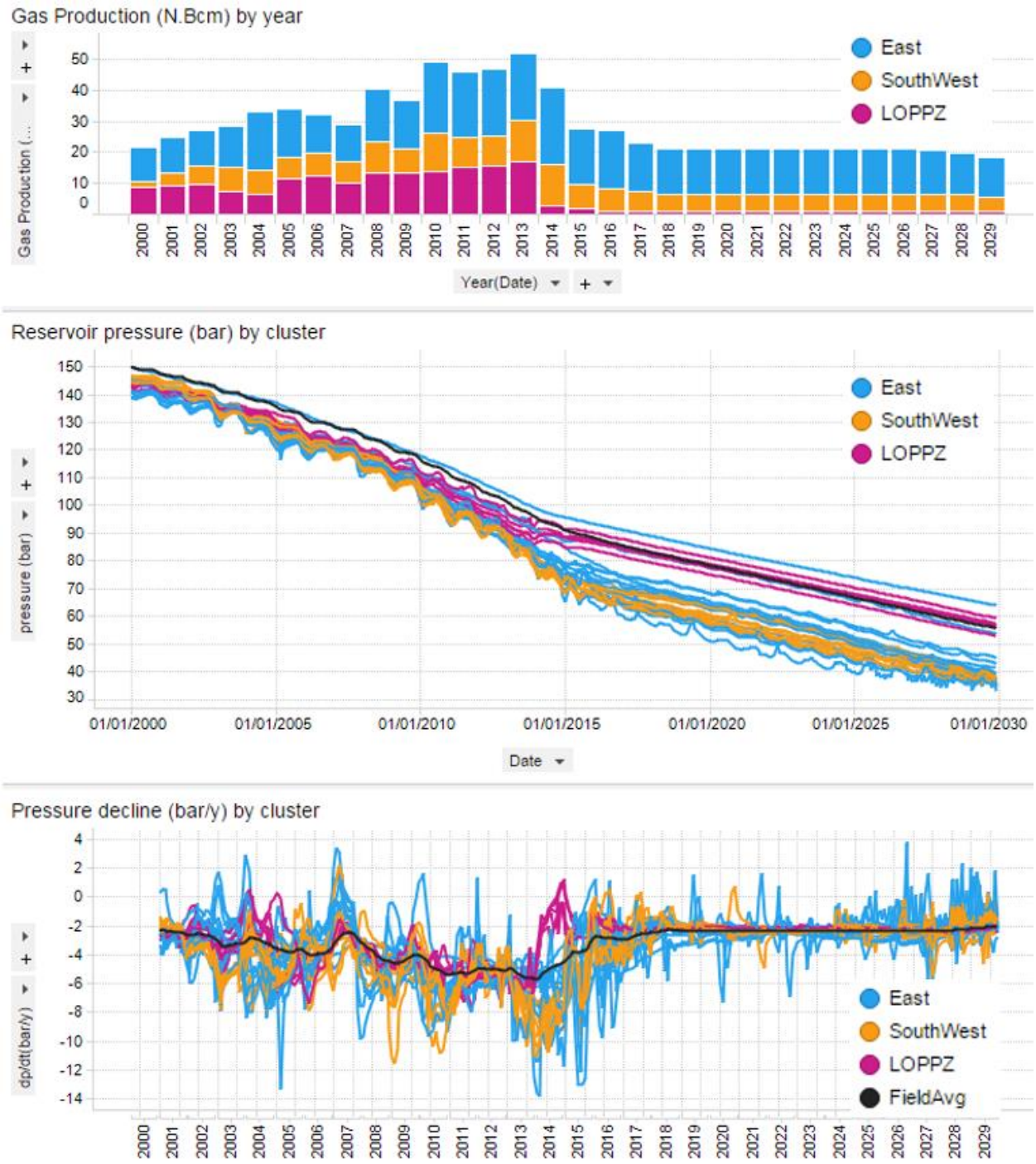


Figure 1-5: Overview of simulated reservoir pressure response to production rates, for both history and forecast (BP2017 NFA). Both the production rates and pressures are colored by the regulatory regions as outlined in Chapter 1.4.1 (in order not to make the graphs unnecessary busy, Eemskanaal is omitted). Individual pressures trends are plotted for each production cluster.

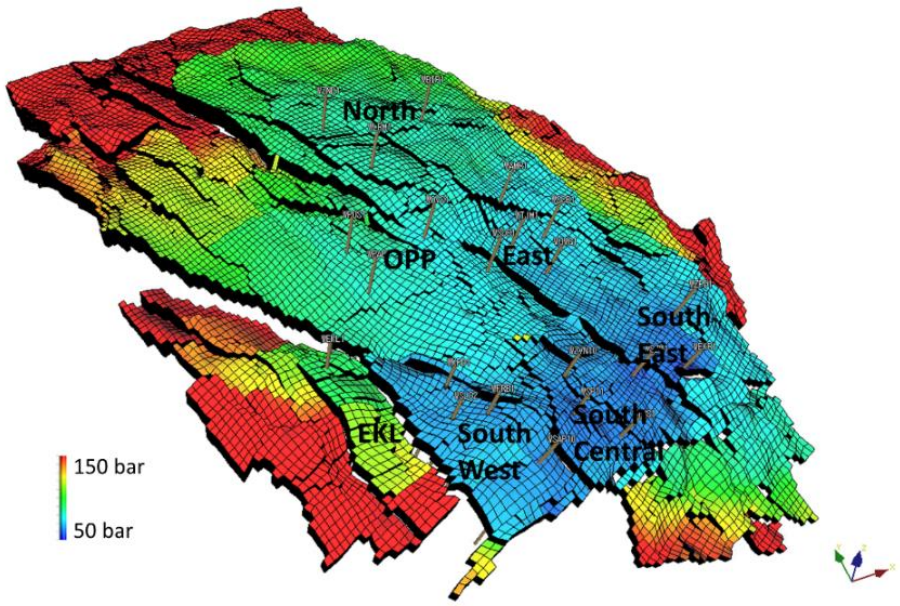


Figure 1-6: Areal distribution of pressure in the Groningen field as of 2017 (color scale clipped at 150 bar)

Region	Clusters				
North	BIR	ZND	LRM		
East	AMR	SCB	TJM	OWG	SDB
OPP-E	OVS	POS	PAU	EKL	
South-East	EKR	SZW	ZPD		
South-Central	SPI	ZVN	TUS	SAP	
South-West	SLO	FRB	KPD		

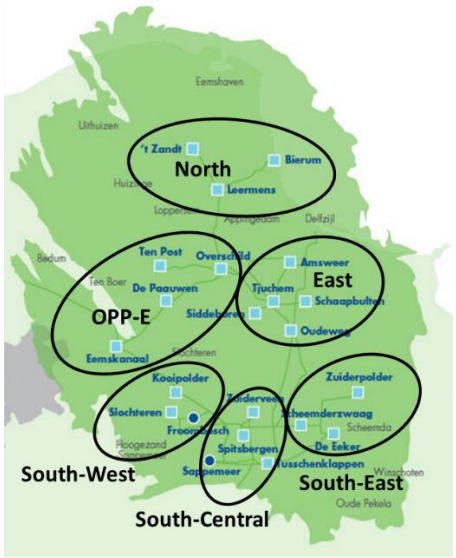


Figure 1-7: Production clusters with respect to the various areas of the field.

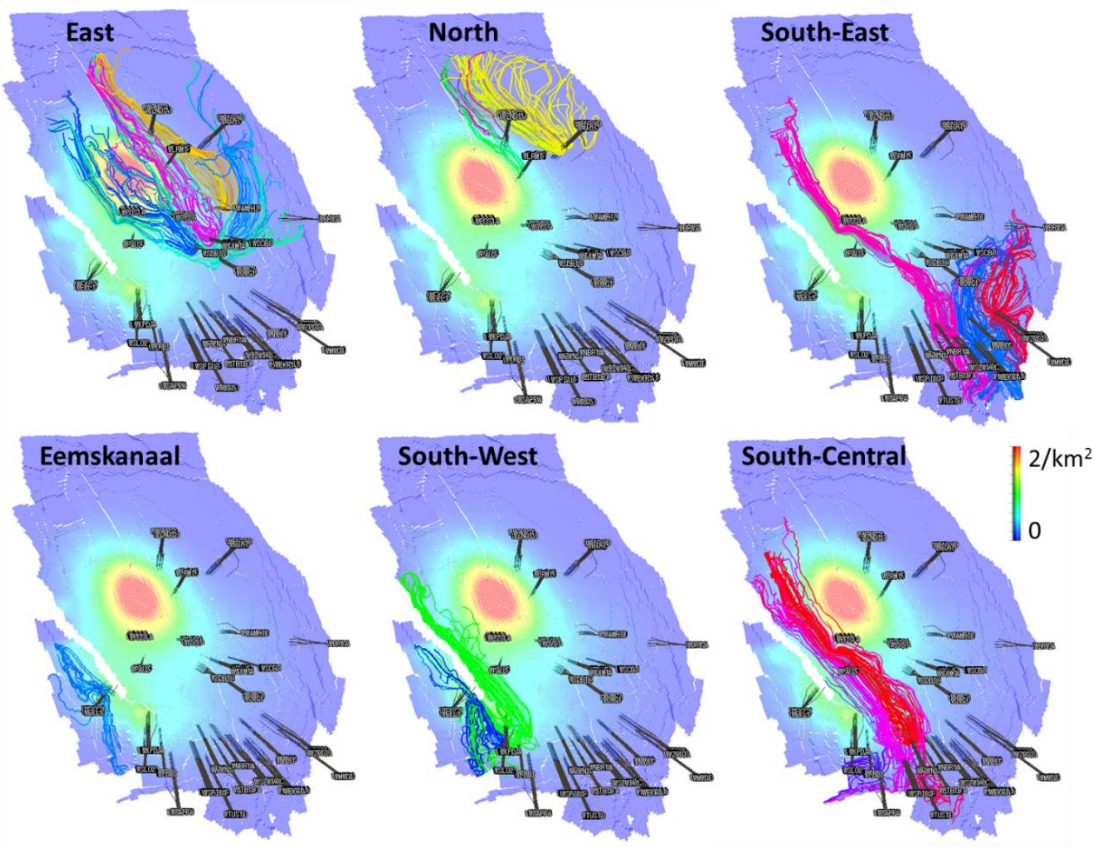


Figure 1-8: Streamlines as per current production distribution (1/1/2018). The streamlines are colored by arriving producer, and dedicated visualizations are provided by areal group of production clusters: East, North, South-East, Eemskanaal, South-West and South-Central. Given their very low flowrates, the Loppersum clusters are omitted from this graph. The underlying property is (modelled) cumulative earthquake density as per 2018.

1.5.3 Geological impact on gas flow

From References [5, 6, 7], it is clear that seismicity has a relation to zones of faulting in the reservoir. Some of the major faults in the reservoir are clearly visible from Figure 1-9. Although the entire field is in pressure communication, as is clear from Figure 1-6 and Figure 1-8, these faults do act as pressure baffles that will impact the spatial pressure depletion and compaction patterns that result from re-distributing production offtake. Hence, they will subsequently impact seismicity.

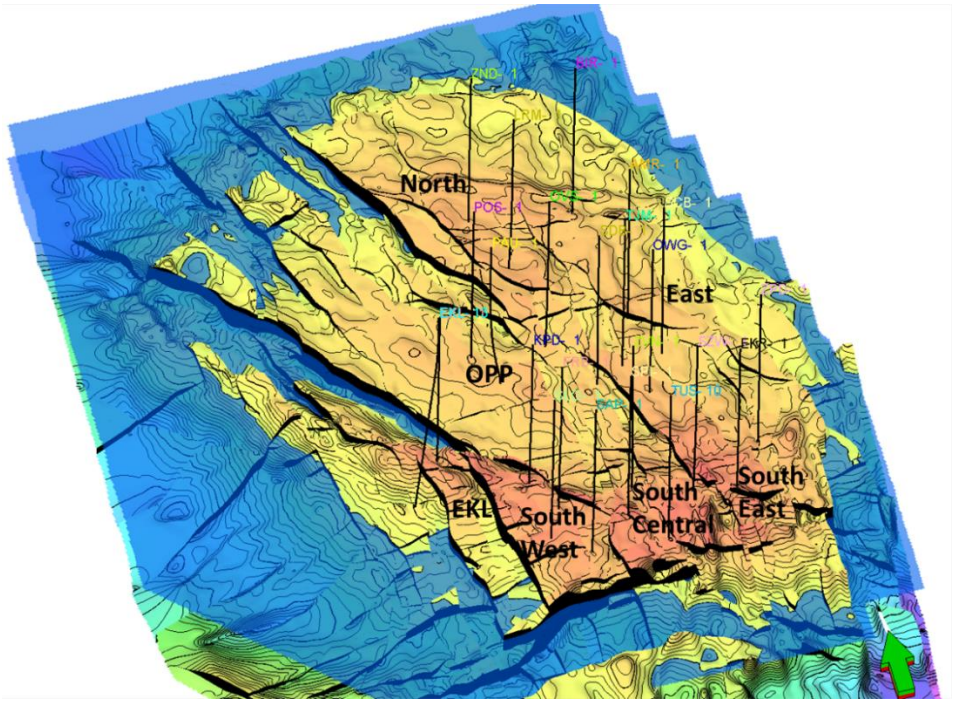


Figure 1-9: Geological map of the Groningen field showing top surface height. The blue plane reflects the original gas-water contact.

The streamlines shown in Figure 1-8 reflect the pressure trend that currently exists in the reservoir, whereby the structural influence can be clearly observed. Most of the streamlines follow a NW-SE trend that align with the major direction faulting. Although less apparent, this trend is also visible when plotting streamlines from 2011, when production distribution was still managed such as to establish a single reservoir pressure across the entire field, Figure 1-10. Specifically, the major bounding faults (and in-between graben) separating the north-western and north-eastern part of the field can be clearly seen to restrict connectivity in the East-West direction.

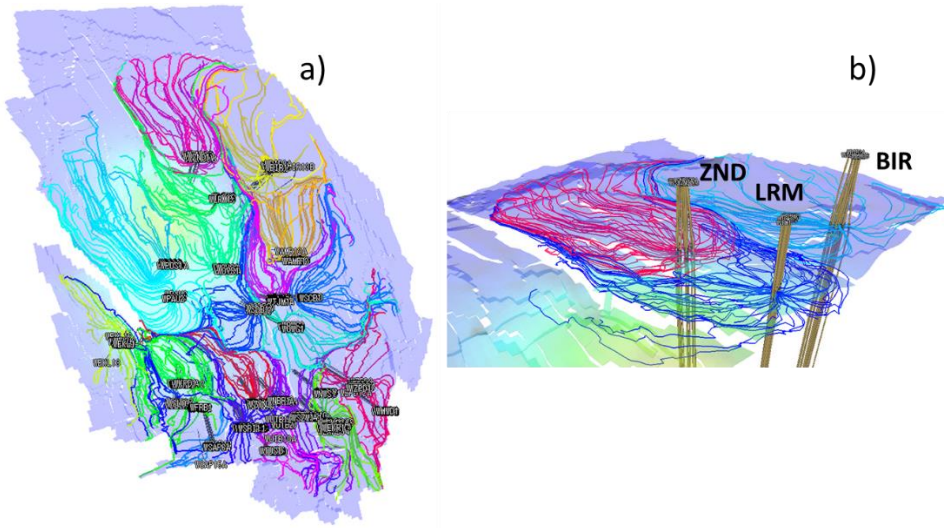


Figure 1-10: Streamlines as per production distribution in 1-1-2011 for the entire field in map view (a) and specifically for the North area in oblique view (b). Streamlines are colored by arriving producer and the underlying property is modelled earthquake density as per 2011.

2 Modelling approach

2.1 Cause and effect chain

On a high level, the cause and effect chain with respect to production induced seismicity is modelled as follows:

⇨ Production from clusters

⇨ Pressure depletion

⇨ Compaction

⇨ Event Rate (Earthquakes)

⇨ Hazard

⇨ Risk

2.2 Model implementation

In the regular Hazard and Risk Assessments that NAM issues, the steps in the cause and effect chain up to "Pressure depletion" are calculated in the dynamic reservoir simulator (Mores, part of the Dynamo tool suite[5]), as described in Reference [8]. The subsequent calculation steps to predict earthquakes, hazard, and/or risk, are calculated in the seismic Hazard and Risk Assessment (HRA) model as described in References [9], [6] and [7].

A model representation of the cause and effect chain can be implemented within a control loop, which allows for a model driven optimisation. This production optimisation study consists of two parallel alternative work streams:

1. **Stand-alone Mores seismicity proxy**

   In addition to pressure depletion, two subsequent steps in the chain of effects were implemented within the dynamic simulator: Compaction and Event Rate. This set-up allows for a stand-alone optimisation within Dynamo/Mores. Note that this (single deterministic) implementation of the Event Rate was only a proxy for the full analysis as done by the HRA model.

2. **Mores-HRA coupling**

   For a more robust optimisation, an integral coupling of Mores with the HRA model was also pursued. In 2016, the runtime for the reservoir simulator in combination with the hazard and risk assessment tool was deemed too long to allow for optimisation based on a combined tool. A software development team was constituted to speed up the runtime of the hazard and risk assessment tool. This team has been able to reduce the runtime of the hazard and risk assessment tool very significantly, allowing development of a practical optimisation tool based on integrating

[5] Shell proprietary software

the hazard and risk tool with the reservoir simulator. This was done by optimisation of the code and exploiting benefits of further parallelisation.

Work stream 2 was seen as working towards a longer-term goal, while Work stream 1 was the main activity delivering the milestones in the Instemmingsbesluit in a robust manner. In the Wijzigingsbesluit, the deadlines for the Draft (Concept) Report and for the Final Report were shifted by 1 month, which also made optimisation using Work stream 2 possible.

A high-level overview of the respective work streams and their calculation steps is given in Figure 2-1. For both work streams the reservoir pressures are calculated in Mores for a given distribution of gas offtake across the production clusters. The optimisation is run within the Dynamo tool suite which includes optimisation algorithms. The HRA model is run in Python, using a pressure grid from Mores as an input, providing estimates of events, hazard and risk as output. This output is subsequently imported back into Dynamo and used during the optimisation.

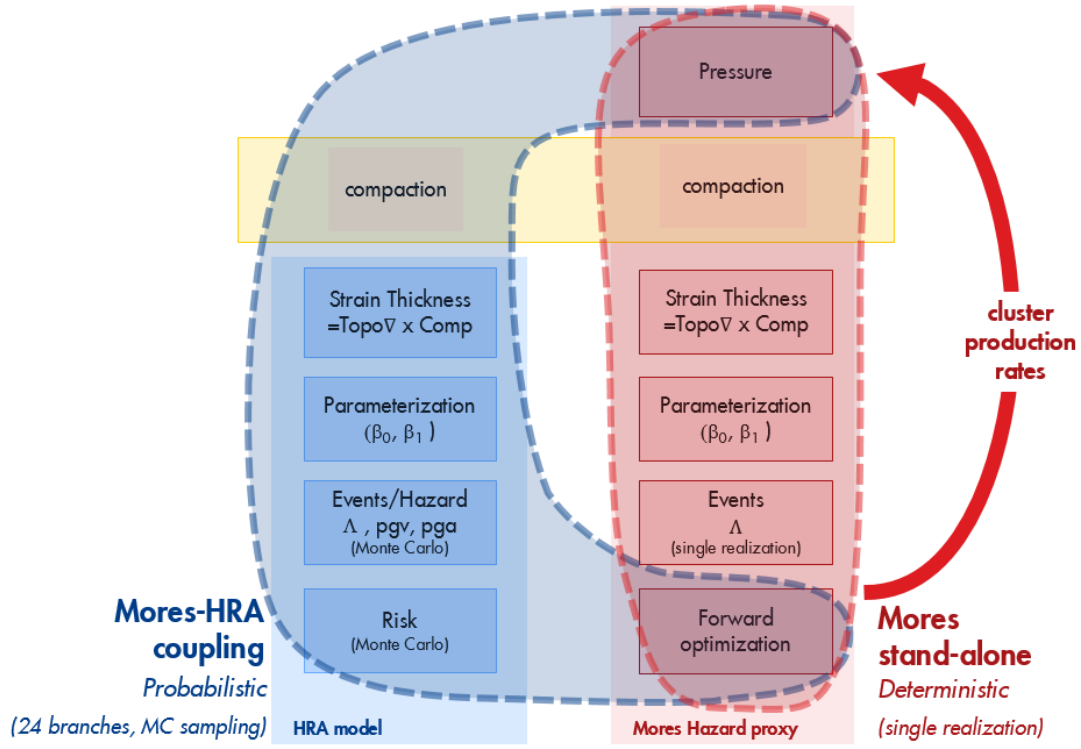


Figure 2-1: Model workflow for both workstreams, the stand-alone Mores proxy (in red) and the Mores-HRA coupling (in blue)

2.3 Model differences

There is a fundamental difference in approach between the traditional dynamic reservoir simulation model and the Hazard and Risk model.

2.3.1 Mores model

Traditional dynamic reservoir simulation modelling is setup around a physical process that is well understood: fluid dynamics in a porous medium. Its foundations lie in global academic research that goes back for over more than half of a century, including extensive laboratory testing.

The part of the dynamic reservoir model that describes the reservoir (the static model) combines a multitude of inputs that all have different resolutions.

- Core data: down to grain size (μm-mm)
- Log data: 1ft vertical resolution, couple of feet lateral resolution (penetration within the reservoir)
- Seismic data: 30m vertical resolution, 50m lateral resolution
- Geological understanding of clastic environments and the Groningen depositional environment in particular

Each of these inputs have associated uncertainties, and the parameter values are typically quoted as a range, with a best estimate value and an uncertainty band. The static model is represented using a dynamic reservoir simulation grid of a certain gridblock resolution (roughly 400x400x10m for Groningen specifically). This involves both upscaling (of e.g. well log data), and downscaling (e.g. geological models are generated to interpolate trends between wells). There are lots of unknowns that cannot be fully calibrated (e.g. dedicated fault transmissibility factor for all 1,100+ faults in the Groningen reservoir model).

The common approach in Petroleum Engineering is to generate a multitude of deterministic model representations (typically in a Monte Carlo type approach) which each fit the input dataset within their respective uncertainty bands. These models are than assessed with respect to the prediction parameter of interest (e.g. Ultimate Recovery for development decisions), and P10/50/90 model realizations are used to represent the solution space.

2.3.2 Hazard and Risk model

The Hazard and Risk Assessment for the Groningen field provides a model based representation of complex physical processes using a probabilistic approach based on a limited set of earthquake data. Based on a rigorous assessment of the remaining uncertainties, this allows a probabilistic forecast to be prepared.

The data types to which the Hazard model is calibrated includes subsidence data and seismic event data. Based on an evaluation of model quality against these available datasets, the appropriate level of smoothing/model resolution was established for the HRA model.

Subsidence

Surface subsidence measurements are used to derive an estimate of reservoir compaction. These measurements typically have a temporal resolution of 5 years (leveling surveys). When converting from surface to reservoir, the overburden acts as a low pass filter, limiting the spatial resolution of the

compaction estimate to about 3 km. It was established that fine scale compressibility variation is fairly insensitive to the overall observed subsidence, see Appendix 2.

Events

The seismic dataset for the Groningen field contains 271 events (of magnitude greater than 1.5) as of 31/12/2016, see Figure 2-2, which is rather limited for bringing temporal/spatial resolution into the model (section 2.4). The available number of events is a function of the magnitude of completeness: the earthquake magnitude that can be detected at any place in the field. As a function of the surveillance metering equipment, the magnitude of completeness was M=1.5 from 1995 onwards, Reference [4].

Note that the 2015 upgrade of the shallow geophone network operated by KNMI has increased the Magnitude of Completeness to M=0.5. This results in principle in 10 times more data available for the model calibration (based on a Gutenberg-Richter type logarithmic description of earthquake magnitudes). On top of the limited data resolution in time, there is also a spatial uncertainty in the epicenter location estimate. For the earthquakes recorded using the geophone network prior to 2015, this was estimated to be some 500 m. However, with the denser network of geophone stations since 2015 and the development of new methods to establish the earthquake hypocenter based on full waveform inversion, the uncertainty in the lateral location of the hypocenter is reduced to some 100 m.

Given these spatial and temporal data resolution restrictions, the HRA model was found to have a best fit at a lateral resolution of 5 km, Reference [9]. This resolution yields the highest model quality (falsifiable model), thus avoiding either an overdetermined or underdetermined model (see section 2.4). It is plausible that in time, with the expansion of the calibration dataset, higher resolution models can be fitted. Due to the improved magnitude of completeness since 2015, and accumulating additional data as time progresses, the lateral resolution of the HRA model can likely be reduced to about 3km (despite the additional uncertainty with respect to b-value). Further studies are planned to achieve a higher resolution. Note that an implication of the limited resolution in the HRA model is that pressure effects in the reservoir at a smaller length scale (e.g. pressure differences over faults) will disappear as a result of the smoothing.

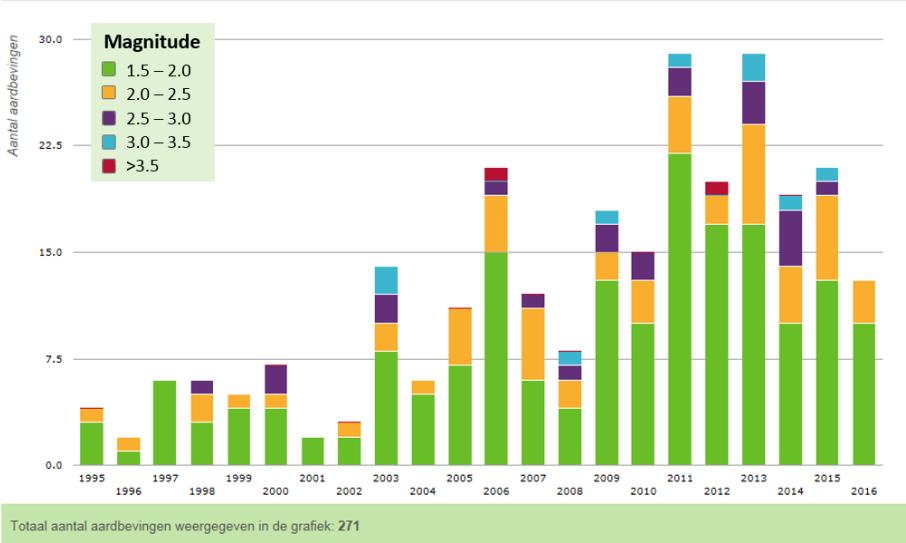


Figure 2-2: Overview of measured earthquakes (M>=1.5)

2.4 Model resolution

Figure 2-3 schematically illustrates the impact of data availability (e.g. sampling frequency of the physical system) on model resolution. A low sampling frequency can potentially result in an underdetermined model. It is possible that an underdetermined model may not reflect the full optimisation potential within the solution space that is span by the available controls. In other words, when smoothing too much potential optimisations can be overlooked.

Given the restricted calibration data set, the accessible spatial event resolution in a data driven seismological model is about 3-5 km (section 2.3.2). It is likely that the actual physical Groningen system operates on a higher resolution than what can currently be resolved by the HRA model. Hence, within the solution space there may be scope for optimisation that cannot be reflected by the HRA model.

On the other hand, optimisation using a potentially overdetermined model would leads to an overestimate in the ability to resolve things. An overdetermined system would likely have a sharper optimum than an underdetermined model (Figure 2-4), but one could be steered towards a (model) optimum which does not necessarily reflect the true optimum of the physical system.

It is recommended to establish whether the conclusions from the optimisation study are sensitive to the resolution of the HRA model. A comparison between the optimisation results from both work streams serves as a good first step, in light of the fact that the Mores stand-alone seismicity proxy retains a higher resolution in terms of reservoir pressure (involves less smoothing) than the HRA model.

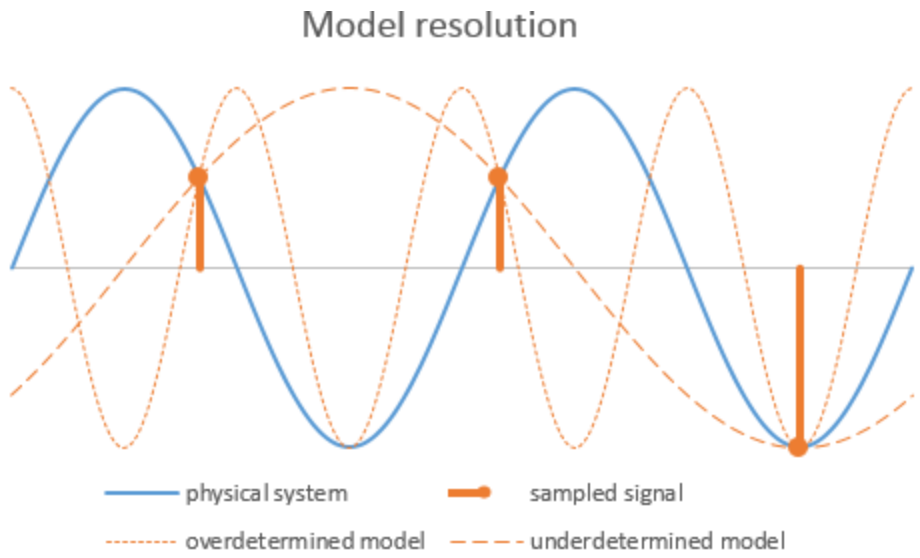


Figure 2-3: Schematic illustration of the impact of sampling frequency on model resolution

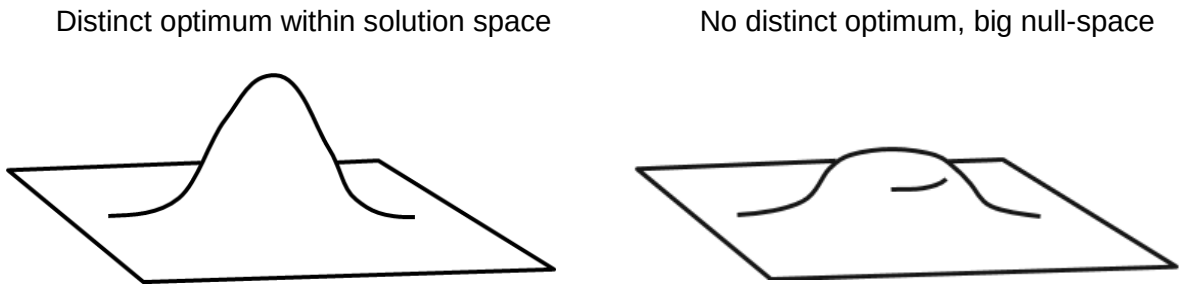


Figure 2-4: Schematic illustration of optimisation

2.5 Modelling restrictions

The current understanding on the Groningen seismicity and subsequent hazard and risk is at the forefront of scientific research, Reference [10]. However, the models can only be calibrated to observed data. With time, more data will come available for calibration. For instance, the signal-to-noise levels may come down for the GPS or InSAR data. This would allow for calibration of the subsidence modelling on a higher temporal resolution, potentially seeing a crisper response to any changes in compaction of the reservoir. As the research evolves, better understanding can develop and the models will be able to better capture the underlying physical processes.

The HRA model is typically used to deliver multi-year estimates, with its inputs (namely the pressure grid) resolved on a year-scale. Hence intra-year effects are not captured. Any impact from rate changes and subsequent pressure transients is not captured. Neither is the possibility of creep-type processes that may reduce the accumulated seismic moment through non-seismic slip.

3 Implementation of optimisation

3.1 Optimisation algorithm

Within the Dynamo tool suite, optimisation capability is implemented. The SPMI algorithm was selected (Simultaneous Perturbation and Multivariate Interpolation), which is further described in **Error! Reference source not found.** and References [11] and [12].

3.2 Objective function

For the Mores-HRA coupled tool the objective functions are based on:

- Nuisance: Event Count
- Hazard: Maximum PGA (475 y)[6], Maximum PGV (475 y) [7]
- Risk: Population weighted PGV[8]

The Event Count is a cumulative over the entire field, and as such represents the impact of a certain production scenario on the entire field. The Hazard objectives (maxPGA and maxPGV) are established as the highest recorded value on the hazard map associated with certain a production scenario. As such, these do not provide a response across the entire system, other than that no other place in the field is modelled to experience a hazard that is higher than this value.

The population weighted PGV was chosen as a proxy for total population risk. This is based on analogy with the PAGER effort in USGS (https://earthquake.usgs.gov/data/pager/), which provides a rapid assessment of the impact of earthquakes (in terms of fatalities and economic loss) and estimates the population exposed to different levels of shaking. In the PAGER method, the Modified Mercalli Intensity scale (MMI) is used to represent ground shaking combining both observations and measurements. Where measurements are available PGA and PGV are converted to MMI. The paper by Wald et al. (1999) (Ref. [13]) shows that PGV correlates with a wider range of macro-seismic intensities, whereas PGA saturates at higher levels of MMI. Once the population within different bands of shaking has been calculated, empirical vulnerability functions are used to estimate the losses (Ref. [14]). Furthermore, the paper by Allen et al. (2009) (Ref. [15]) states that earthquake mortality appears to be systematically linked to the population exposed to severe ground shaking (MMI VIII+, which would be PGV values above 30 cm/s according to Wald et al. (1999)). The population weighted PGV again reflects a response of the entire system, it is calculated by multiplying each cell of the hazard map with the population in that location.

3.3 Controls

In principle, each production cluster can be used as an independent control, for which the optimisation algorithm can set production targets. For computational limitations, multiple production clusters can be lumped into a single control.

[6] PGA hazard is generated for the 475 year return period for all surface locations.
[7] PGV hazard is generated for the 475 year return period for all surface locations.
[8] A weighted average PGV is calculated using the population at each location as the weight

3.4 Boundary conditions

The following boundary conditions were imposed for the optimisation:

- Optimise for a 5 year period (2018-2022)
- Fixed production target for each control (i.e. volume constraint assigned to a grouping of 1 or more clusters) for the optimisation period.
- The optimiser is allowed full freedom in the requested production form each control, clusters are allowed to produce anything between 0 and its full capacity.
- The optimiser is allowed to vary all controls in order to minimize the objective, including those of LOPPZ clusters like 't Zandt.
- 21.6 Bcm annual cap, with a flat production profile of 1.8 Bcm/month
- The maximum cluster capacity is imposed by means of a simple fixed THP constraint, providing a first order reflection of the currently installed compressor capacities.
- There are no additional constraints imposed to reflect the surface facilities (e.g. pressure losses in the Groningen pipeline ring)
- Simple cluster uptime assumption of 90% to reflect planned and un-planned shut-downs.

Given the 5-year window (2018 – 2022) for the optimisation and the 21.6 Bcm field cap, there is no requirement for re-staging of the compressors within this optimisation window. Note that re-staging of production clusters (from 1st to 2nd stage compression) would introduce a multitude of degrees of freedom (i.e. sequence, scope, timing), and significantly complicate the optimisation process.

The impact of these simplifying assumptions made for the optimisation is assessed in Chapter 0, where different scenarios are tested with an integrated production system model (IPSM) that includes both surface and subsurface constraints, and also for a longer time window, including compressor re-staging.

3.5 Subsurface Model

For the optimisation study, the Mores model vintage V4 was used, Reference [8], as was used for the June Hazard update. In Figure 3-1 some typical comparisons between modelled and measured pressures are shown. The production clusters are typically abbreviated by three letters (e.g. LRM for Leermens), these abbreviations are given in Appendix A.

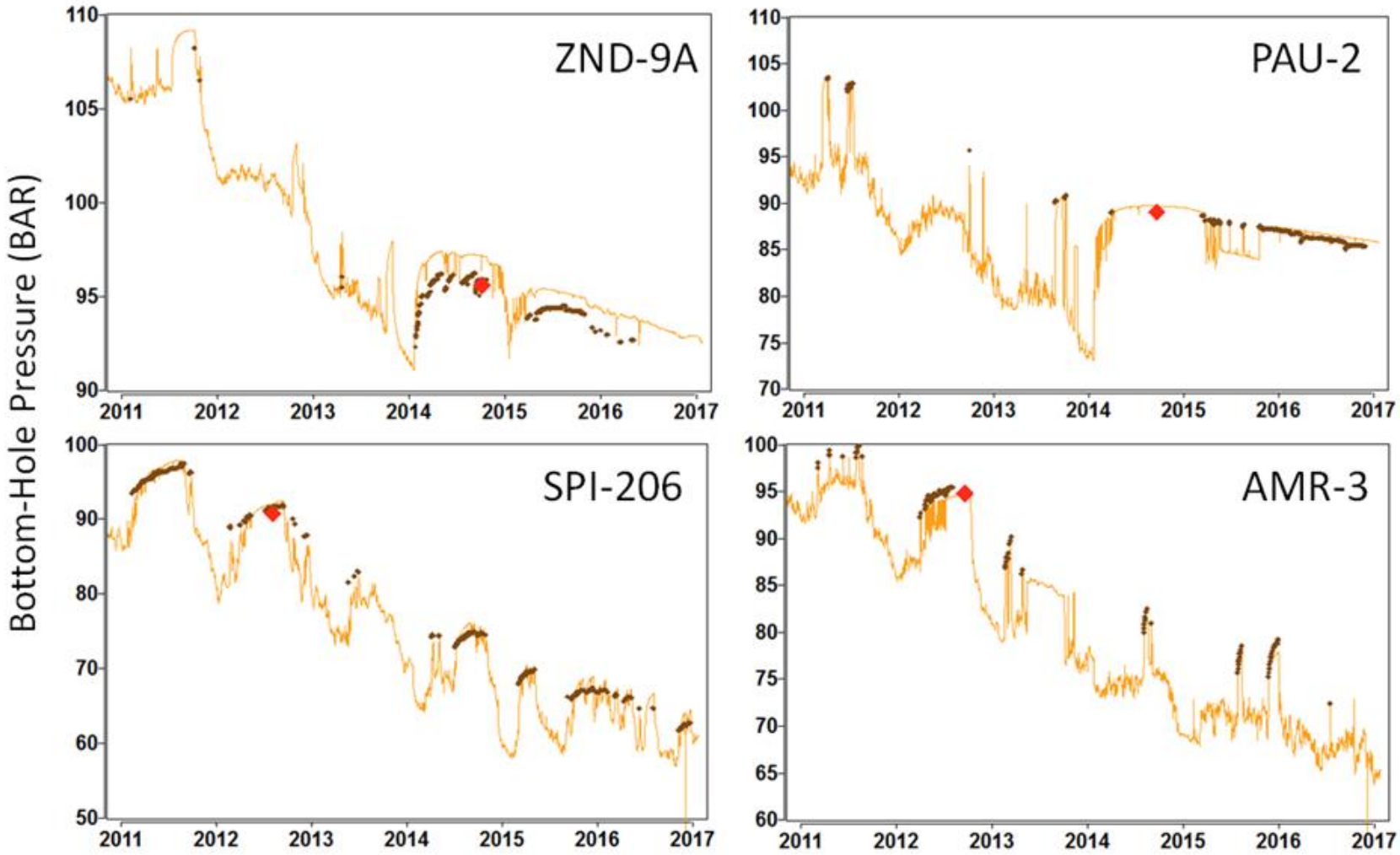


Figure 3-1: Comparison of modelled bottom-hole pressure (lines) to measured bottom-hole data (SPTG, red points) and measured closed-in surface data converted to bottom-hole conditions (brown points) for selected wells.

4 Optimisation – Mores stand-alone

4.1 History match seismicity model

The Mores dynamic reservoir model was upgraded to calculate strain thickness and seismicity in time (single, deterministic realisation), Reference [3]. As can be observed in Figure 4-1, this Mores stand-alone model generates a field-wide seismicity estimate that is reasonably close to historic data.

The areal history matching requires striking the right balance between statistical significance and areal resolution. The historic earthquakes are discrete events. Representing those in a model of seismicity as a continuous function of compaction in time, and acknowledging the probabilistic nature of the seismicity, imposes a need for some form of areal upscaling, achieved through the use of a Gaussian smoothing kernel, Figure 4-2. However, it is important that sufficient areal resolution is preserved in order to allow for a meaningful areal optimisation of production offtake. Figure 4-3 shows an areal breakdown of the match in event rate over the field. Figure 4-4 allows for a visual comparison of the modelled areal response with respect to the actual annual observations over the period 2013-2016.

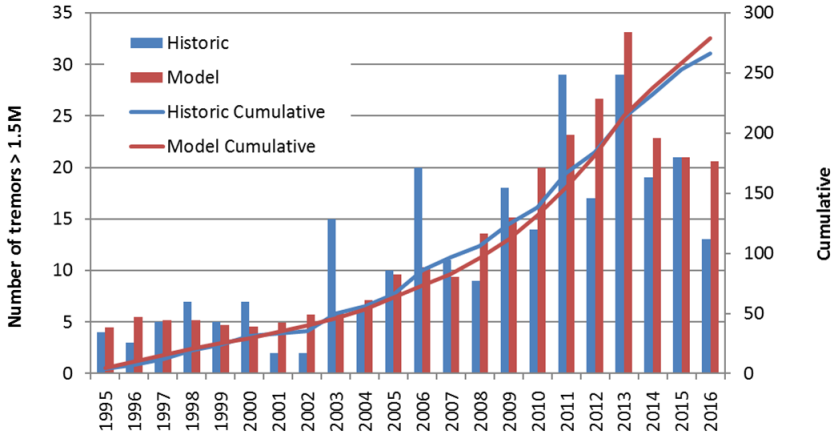


Figure 4-1: Historic and model predicted earthquakes greater than magnitude 1.5 across the Groningen field. Results are based on Mores stand-alone model.

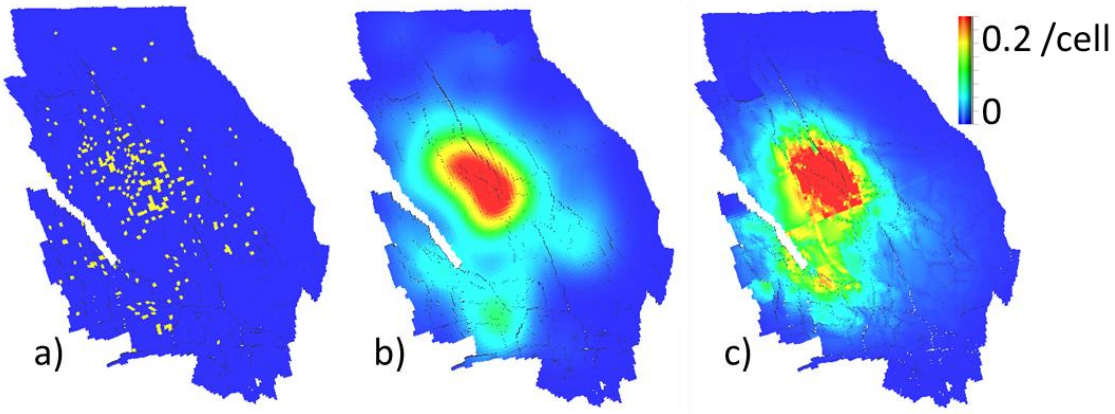


Figure 4-2: (a) Historic locations of earthquakes of magnitude greater than 1.5 at the end of 2016. (b) Historic earthquakes smoothed using a Gaussian filter kernel with standard deviation of 2 km. (c) History matching of cumulative earthquakes using a thin-sheet strain thickness model combined with the dynamic reservoir model (Mores stand-alone model).

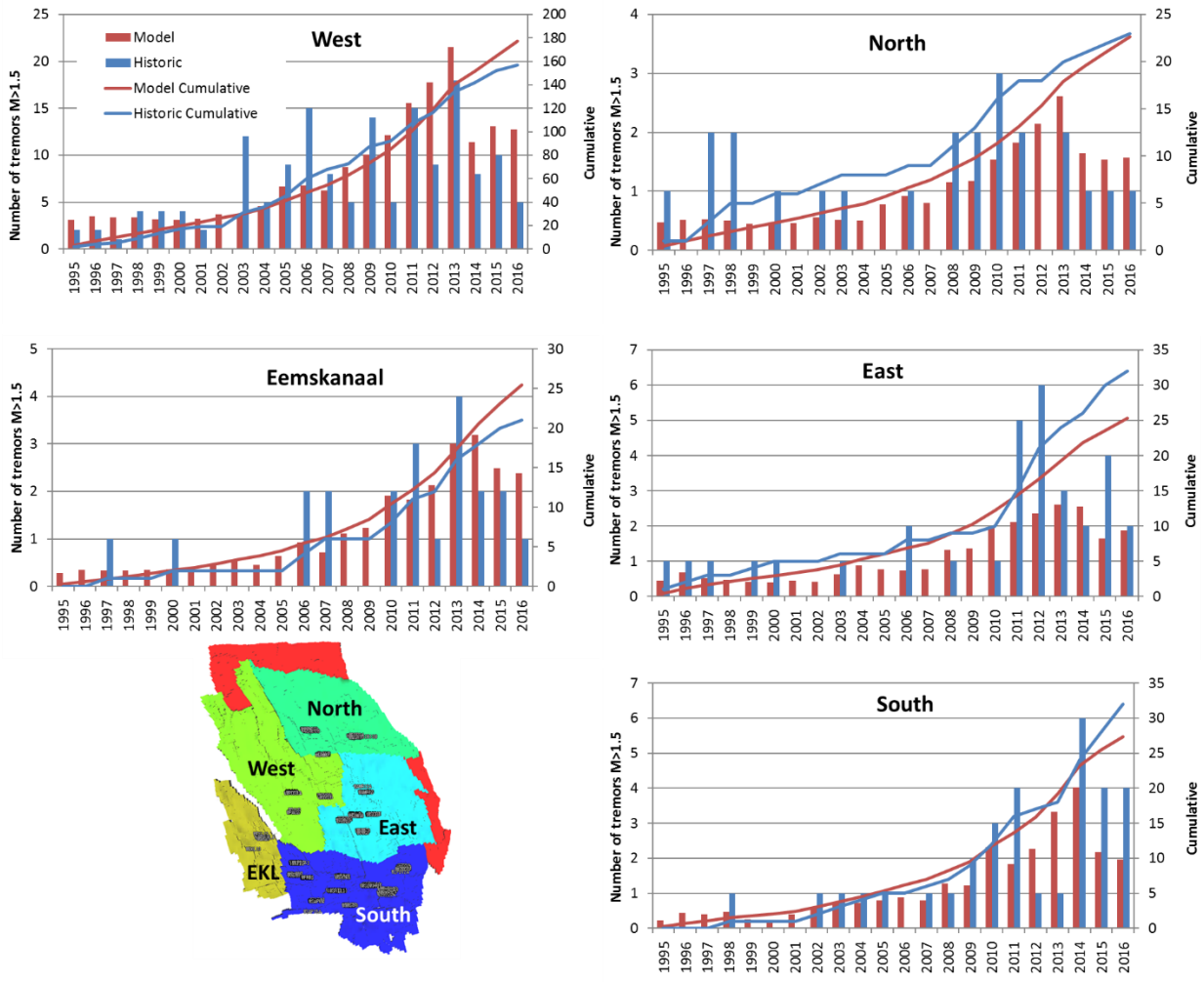


Figure 4-3: Historic and model forecasted earthquakes greater than magnitude 1.5 grouped by regions across the Groningen field. Results are based on Mores stand-alone model.

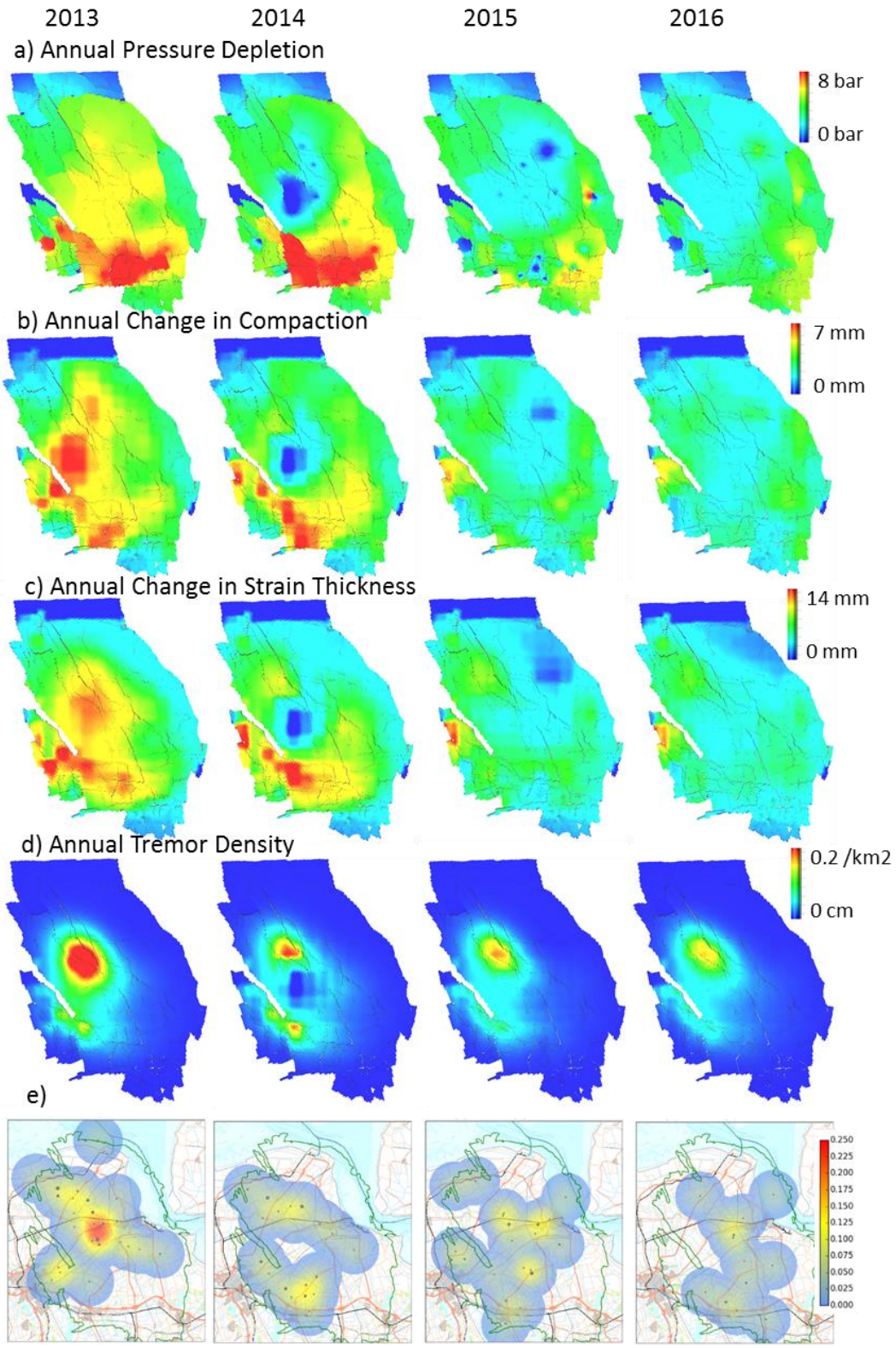


Figure 4-4: Annual changes in properties (2013-16) as estimated by Mores stand-alone model. Actual Earthquake Density maps are shown in e) for M >= 1.5 in units of tremor/km^2/year.

4.2 Model performance testing/Predictive capability

The forecasting capability of the model was assessed by a predictive test (i.e. calibrating the model based on part of the history, and predicting the remainder of the history). The changes in field production following the LOPPZ restrictions provide an opportunity to test the ability of the model to forecast seismic event rate. The main changes in the production (both field offtake and distribution over the field) are:

- A first production cap was imposed on 17/1/2014, constraining the production from the LOPPZ clusters to 3.0 Bcm/y.
- On 14/4/2015 this constraint was further enhanced by reducing the LOPPZ production to stand-by rates for security of supply only (some 1.6 Bcm/y).
- The seasonal swing was reduced from 2015 onwards.

Figure 4-5 shows the impact of these changes on the daily production.

A predictive test was done by calibrating the reservoir model with historical seismicity up to 1/1/2014. Next the response was assessed for the remaining years of seismicity data, which was found to yield a good match, see Figure 4-6. The performance of the model following 2014 is almost equal to that when using the entire history (see Figure 4-1 and Figure 4-3). The corresponding tremor density plot is given in Figure 4-7 and that is also very similar to Figure 4-4 d). The compressibility model that is underlying the activity rate model is also calibrated to data from prior to 2014 (no levelling surveys have been conducted since).

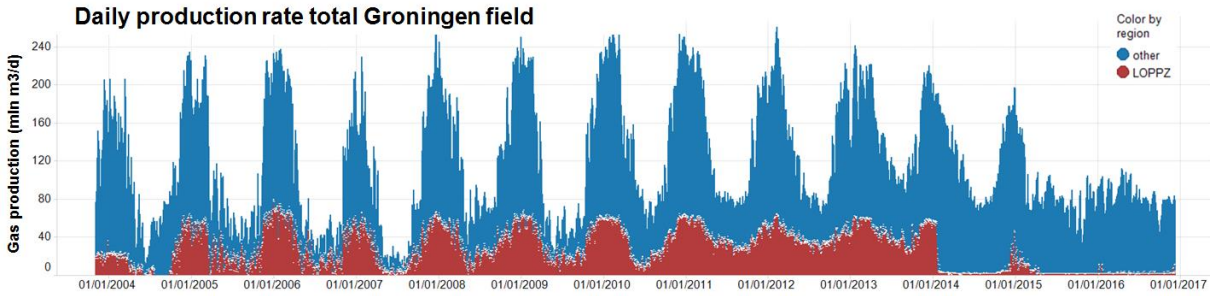


Figure 4-5 Daily production rates for the Groningen field, highlighting the production from the LOPPZ clusters.

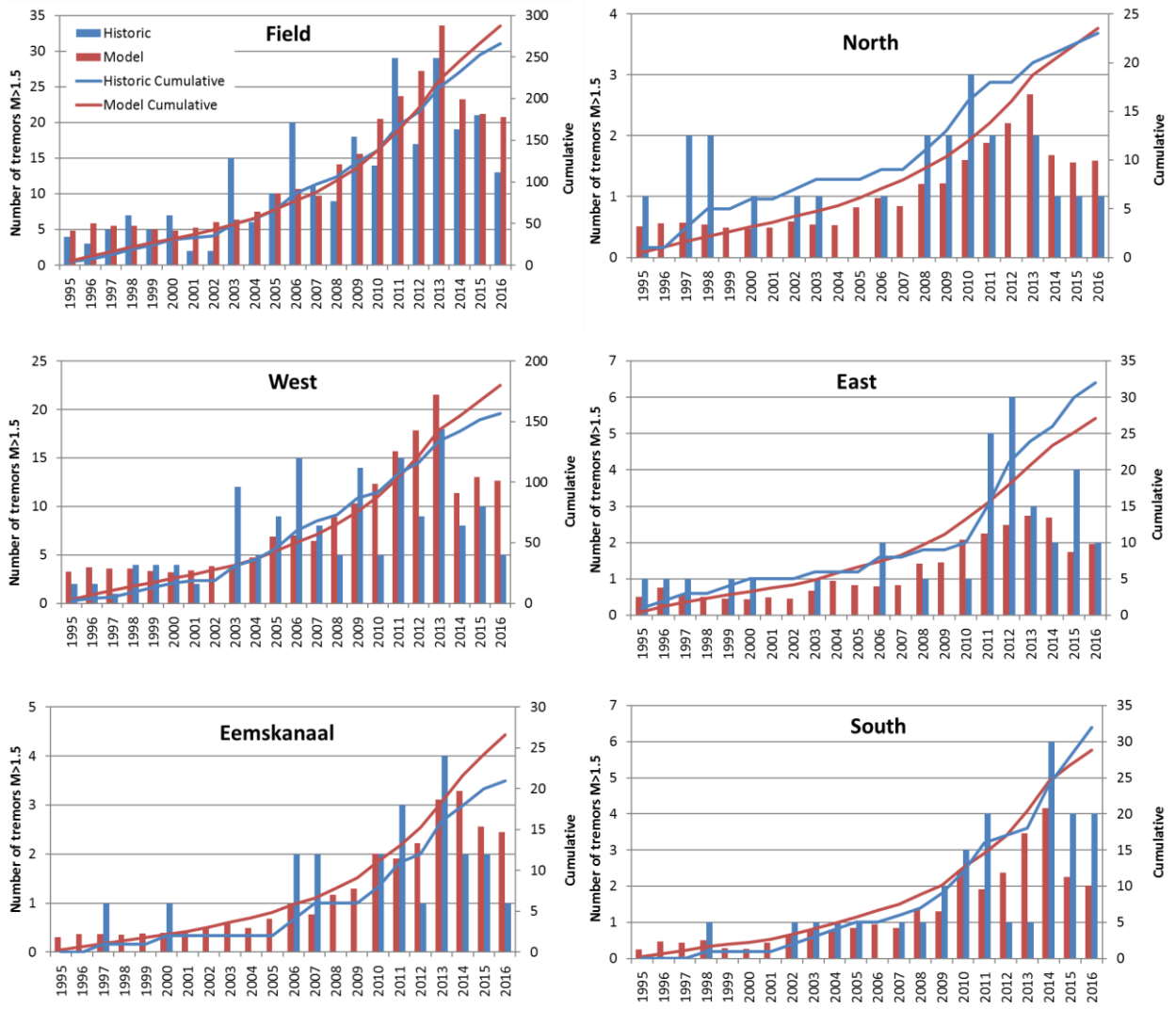


Figure 4-6: History matching seismicity excluding historic data after 1/1/2014 using Mores stand-alone model. Historic and model forecasted earthquakes greater than magnitude 1.5 grouped by regions across the Groningen Field. Regions are the same as in Figure 4-3.

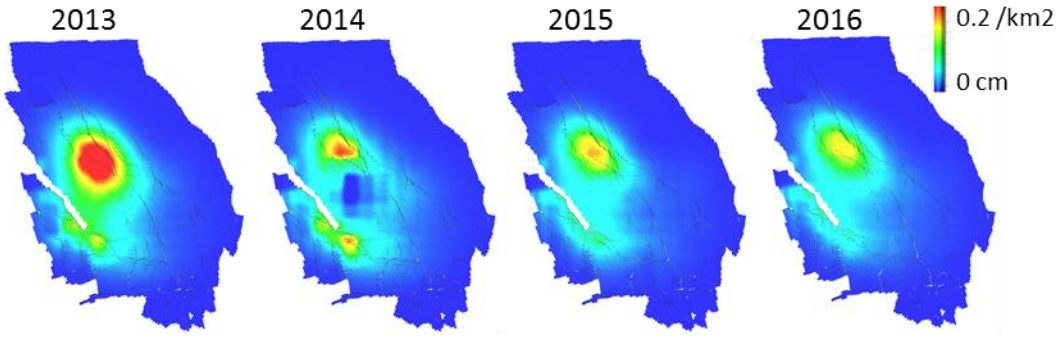


Figure 4-7: Annual change in tremor density (2013-16) as estimated by Mores stand-alone model. History matching seismicity is done excluding historic data after 1/1/2014.

4.3 Model controls

In total, the Mores stand-alone optimisation was assigned 18 independent controls (Figure 4-8). In principle, each one of the 22 production cluster locations can act as an independent control. However, the tie-back locations Sappemeer and Froombosch were lumped. Two sets of two production clusters in close proximity to each other were also lumped: Tjuchem with Schaapbulten, and Zuiderveen with Spitsbergen.

Figure 4-8: Model controls: 18 independent controls. Froombosch is a tie-back to Slochteren, and Sappemeer is a tie-back to Tusschen klappen. Spitsbergen was merged with Zuiderveen, Tjuchem was merged with Schaapbulten.

4.4 Implementation of objective function

Optimisations were run for two objectives:

1. Field wide tremor count, representing a nuisance based metric.
2. Population weighted tremor count, representing a proxy for a risk based metric.

Since the Mores stand-alone model is only estimating tremors, a 2km smoothing kernel was applied to the population map, as a first order reflection of the lateral impact of an earthquake away from its

epicenter, see Figure 4-9. A more rigorous estimation of hazard resulting from ground motion will be included in Chapter 5, when Mores is coupled to the HRA model.

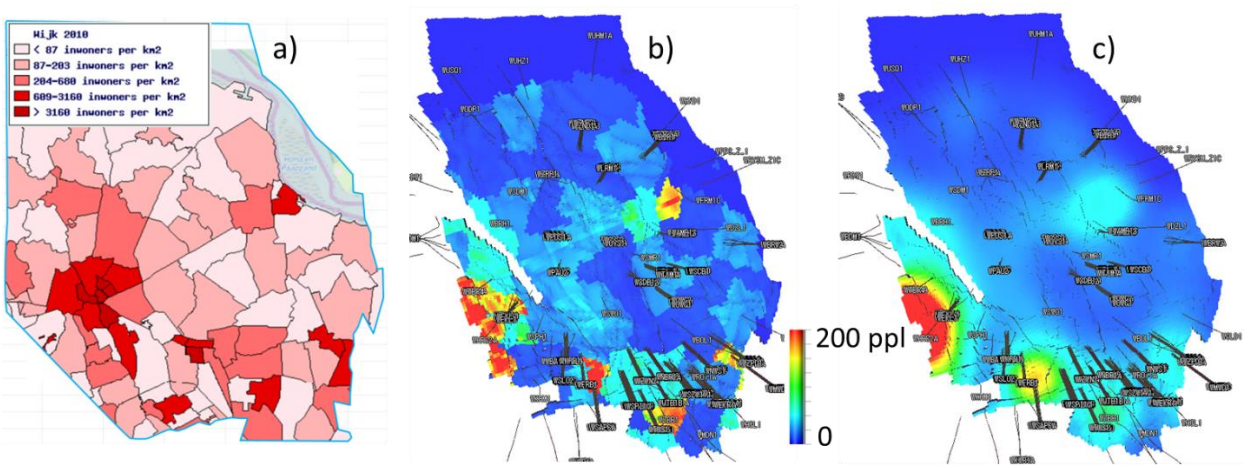


Figure 4-9: a) Population density map (2010 data) from http://kaart.edugis.nl. b) Population mapped onto the grid of the Groningen Field on a per grid cell basis. c) A smoothing kernel is applied to the population map prior to use as a weight factor. In this example a 2km standard deviation is used.

4.5 Optimisation results

The results from the optimisations are given in Table 4-1 with the associated control settings shown in Figure 4-10 and in Figure 4-11 as a bubble map, where the size of the bubbles are scaled to the production fractions of the clusters. The starting point for the optimisation is given by the blue bubbles, and represents the 2017-2021 average production fractions as per NAM's expectation in 2017 (BP 2017). In red are the optimised production fractions. For both optimisation objectives production is removed completely from the West of the field. When including population weighting, some production is shifted from the South towards the North, and the production cluster at 't Zandt is included.

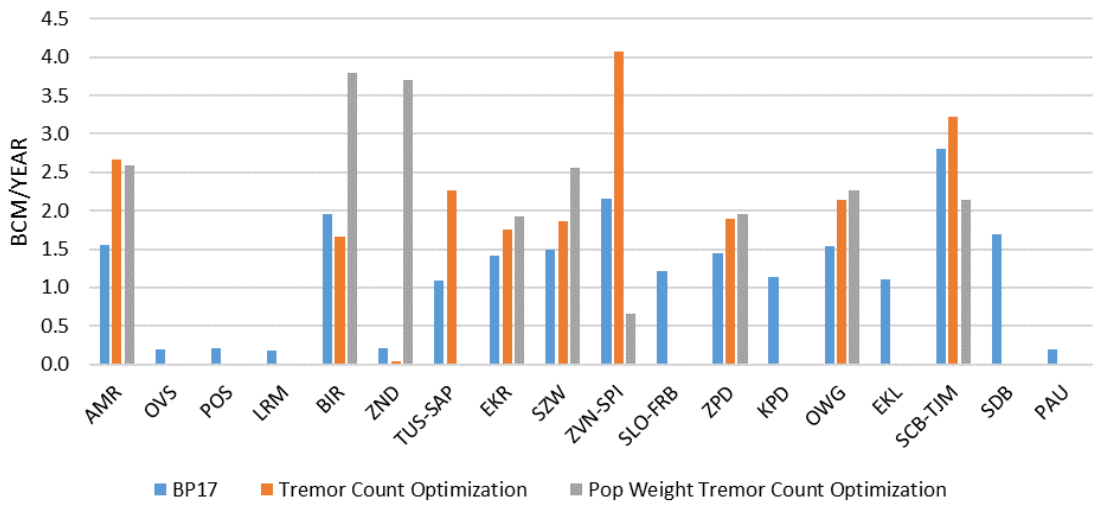


Figure 4-10: Production distribution from the Mores stand-alone optimisation for two optimisation objectives: tremor count and population weighted tremor count.

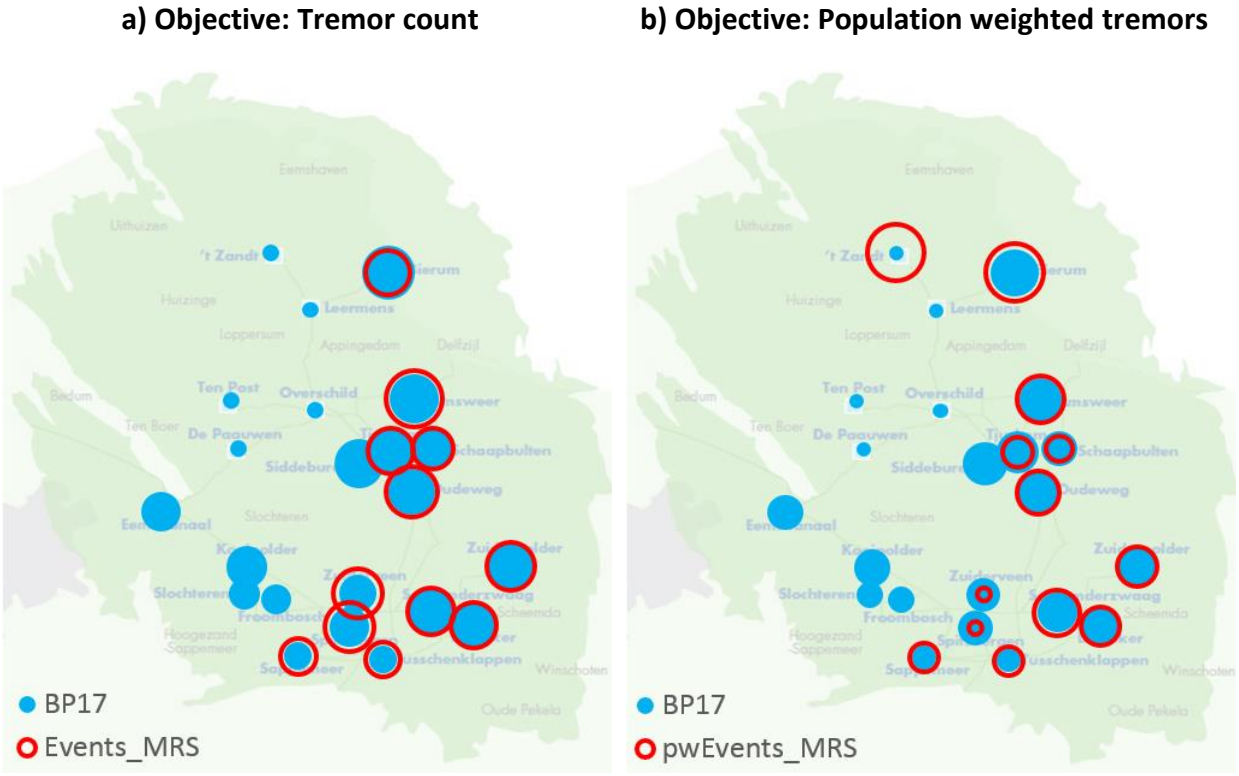


Figure 4-11: Proposed production distribution from the Mores stand-alone optimisation for two optimisation objectives: tremor count (a) and population weighted tremor count (b). The production levels are indicated as a bubble map, with the size of the bubble scaled to the production levels. In blue the current reference case from Business Plan 2017, and in red the optimisation results.

Table 4-1: Optimisation results for MoReS stand-alone in terms of modelled tremors for 5 year period 2017-2021.

Year	BP17 Ref	Tremor Count	PW Tremor Count
2017	20	16	16
2018	21	17	18
2019	22	19	20
2020	24	21	22
2021	25	22	23
SUM	**113**	**94**	**99**
Reduction		***-16%***	***-13%***

5 Optimisation – Mores-HRA coupling

5.1 Model setup

For the optimisation study, the June 2017 vintage of the Hazard and Risk Assessment model vintage was used, which included V4 of the Ground Motion Prediction Equation, Reference [16].

5.1.1 Hazard logic-tree

The HRA model captures the epistemic uncertainty by means of a logic tree. The V4 vintage of the HRA model involves a logic-tree for Hazard with 24 branches, and for Risk of 216 branches. Given the finite calculation capacity the optimisation work was restricted to the Hazard domain only, and population-weighted PGV (pwPGV) was added as a proxy for risk (section 3.2).

The logic tree for the hazard assessment comprises three sets of branches to capture the uncertainty in the different elements, Figure 5-1. The first set covers the uncertainty with respect to Mmax in the seismological model, and the second and third sets capture the uncertainty related to the GMPE (tau and sigma). Each branch of the logic tree represents a scenario, and by combining all scenarios using the weights in the logic tree, the mean hazard map can be calculated.

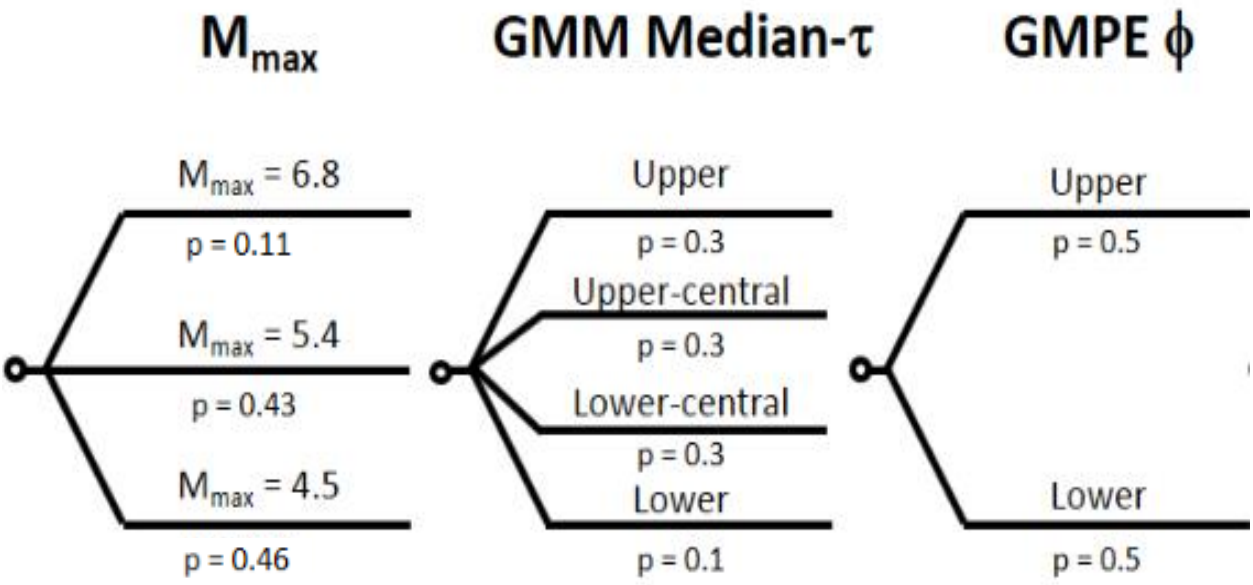


Figure 5-1 Hazard Logic-tree.

5.1.2 Overview of the parameter space

The Mores-HRA coupled model setup allows for optimisation with respect to a multitude of objective functions (or metrics) covering:

- nuisance [Events],
- hazard [maxPGA]/[maxPGV],
- risk as approximated by [pwPGV],

- and additionally a weighted combination of two or more metrics can be used in the optimisation [Hybrid].

The SPMI optimiser starts with a given control setting (initial rates of production clusters), and works its way towards an optimum in a series of iterations (**Error! Reference source not found.**). As partial mitigation of the possibility that the optimisation's starting point plays a significant role in the end-result, three different initial production-distributions were considered:

- Base-case (BP17): in line with the current regulatory steer to minimize production from LOPPZ, and keep a proportional weighting between the East and SouthWest regions.
- Equal rates: equal production rates for each cluster
- Worst case: maximum production from LOPPZ, EKL and the SouthWest

These initial control settings are summarized in Table 5-1, and visualized in Figure 5-3.

Table 5-2 provides an overview of the combined realization table for the objective functions and initial rates.

To ensure that the results are optimised with respect to the entire uncertainty range as captured in the logic tree, a dedicated optimisation was run for each of the 24 logic-tree branches, using the following process:

- First the objective function and initial production distribution is selected
- The following steps are executed in parallel for each branch of the logic tree
    - Given the selected set of objective function and initial rate, a dedicated SPMI optimisation is run
    - An optimum production distribution is obtained
    - The optimisations' end results are each evaluated on the entire logic tree (24 branches) and the value of the logic-tree mean is calculated for each metric using the weights of the various branches

This process setup ensures a consistent Hazard assessment (i.e. taking into account the epistemic uncertainties) and enables a like-for-like comparison of the 24 different optimisation outcomes' performance. Figure 5-2 gives an overview of the process.

In chapter 7 it is described how subsequently the shortlisted distributions are rerun using the subsurface model coupled with a high-fidelity surface network model which reflects the operation limitations of the surface facilities (allowing for more realistic implementation of production distributions) and evaluated using the full logic-tree.

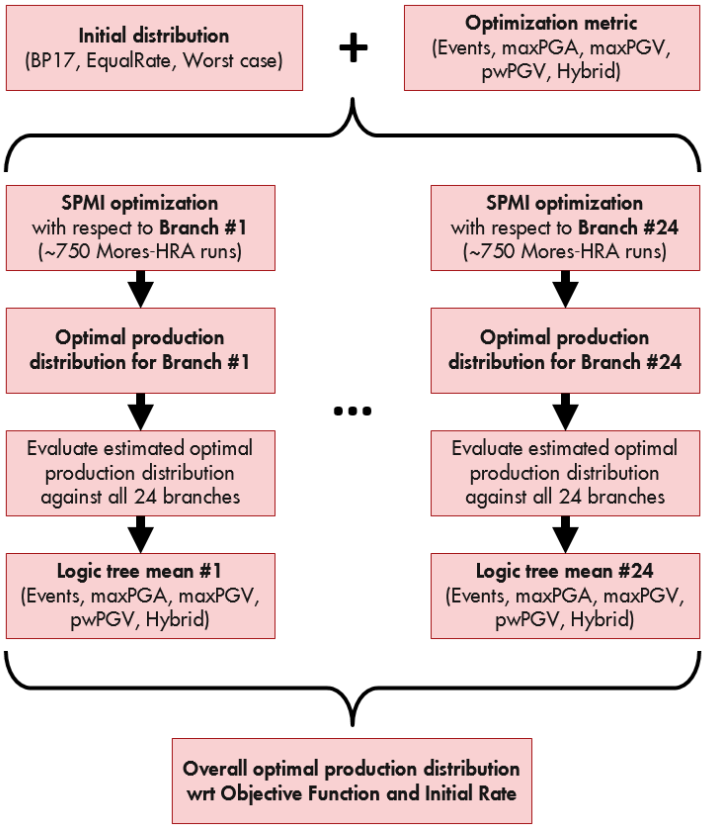


Figure 5-2 Optimisation process-flow for an optimisation metric/initial distribution pair.

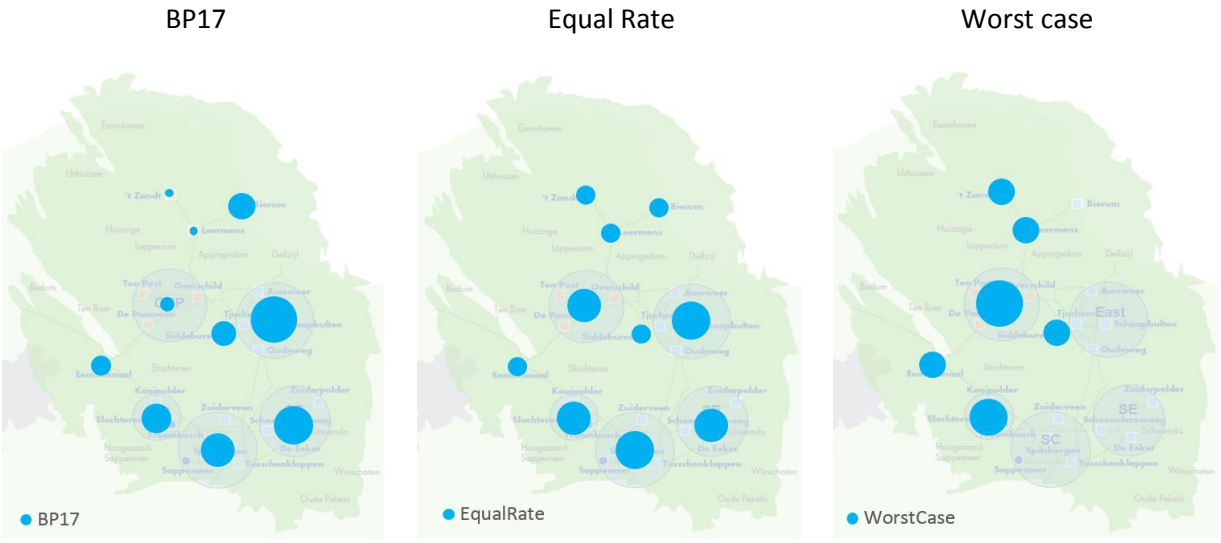


Figure 5-3: Initial production distributions

Table 5-1 Volume fractions and production rates for the initial distributions

control	Fraction			Rate (mln m3/d)		
	BP17	EqualRate	WorstCase	BP17	EqualRate	WorstCase
BIR	0.09	0.05	0	5.4	2.7	0
East	0.27	0.18	0	16.2	10.8	0
EKL	0.05	0.05	0.11	3	2.7	6.6
LRM	0.01	0.05	0.11	0.5	2.7	6.6
OPP	0.03	0.14	0.33	1.7	8	19.7
SCentral	0.15	0.18	0	8.9	10.8	0
SDB	0.08	0.05	0.11	4.6	2.7	6.6
SEast	0.2	0.14	0	12	8	0
SWest	0.11	0.14	0.22	6.5	8	13.1
ZND	0.01	0.05	0.11	0.5	2.7	6.6

Table 5-2: Realization table for the various objective functions versus different initial rate scenarios

		Objective function				
		Nuisance	Hazard		Risk	hybrid
		Event count	Max PGA (475 yr)	Max PGV (475 yr)	Population weighted PGV (pwPGV)	pwPGV+Event count
Initial cluster rates	BP17					
	Equal Rate					
	Worst case					

5.1.3 Model controls

To limit calculation time, the number of model controls was reduced by lumping some production clusters. The following logic was applied:

- Maintain spatial spread of individual controls
- Allow for sequential steps outwards from high seismicity areas.
- Group clusters with similar impact on seismicity metrics.

From analysis of optimisations based on the full set of controls (e.g. Figure 4-11), it was possible to find groupings of clusters that behave similarly given an optimisation metric. The most eastern and south - eastern clusters tend to always produce close to full capacity. The behavior of the south-west and south-central clusters depend on whether population density is part of the metric. Additionally, clusters that are close to the higher seismicity area, e.g. 't Zandt, Leermens or Siddeburen, are considered independent controls.

The total number of controls was reduced to 10, as displayed in Figure 5-4 and summarized in Table 5-3.

Table 5-3: Lumping of production clusters over 10 model controls

	Control	Clusters			
1	East	AMR	SCB	TJM	OWG
2	OPP	OVS	POS	PAU	
3	EKL	EKL			
4	LRM	LRM			
5	BIR	BIR			
6	ZND	ZND			
7	South-West	SLO	FRB	KPD	
8	South-East	EKR	SZW	ZPD	
9	SDB	SDB			
10	South-Central	ZVN	SPI	TUS	SAP

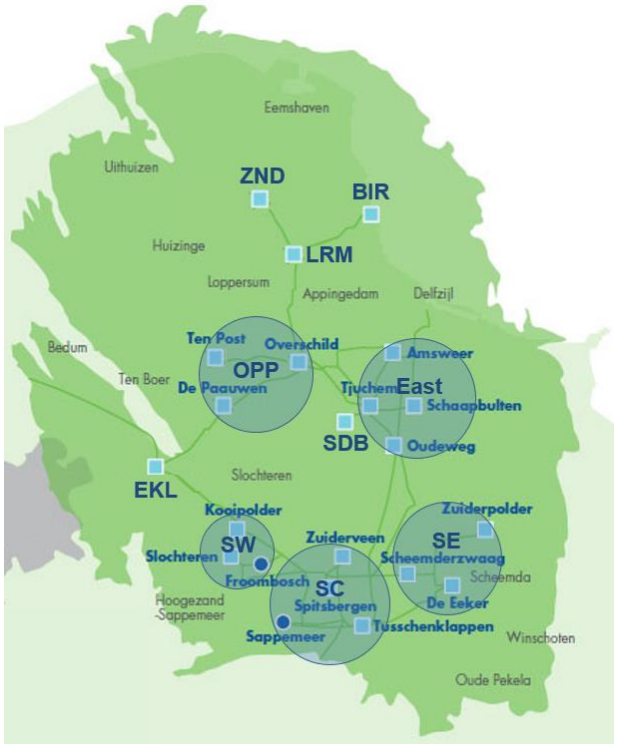


Figure 5-4: Setup for 10 controls, comprising five sets of clusters (OPP, SW, SC, SE, East) and five individual clusters (EKL, SDB, LRM, BIR, ZND)

5.1.4 Overall run-count and computational intensity

In total, over 360 SPMI optimisation runs have been performed:

3 initial distributions × 5 objective functions × 24 HRA branches.

Each SPMI optimisation (see **Error! Reference source not found.**) goes through 15 steps ("iterations"), with each step requiring roughly 35 separate reservoir-simulations ("cases"), each with its own subsequent HRA calculation. This adds up to more than 189,000 reservoir simulations. As explained above, each of the 360+ end-results underwent further evaluation using the full logic-tree, bringing the total number of HRA runs to roughly 200,000.

The runs were performed in-house, in data centres in Houston and Amsterdam over a period of two months. When executed on the available hardware[9], a reservoir-simulation run typically took about 20 minutes, and an HRA run 15 minutes. This brings the total CPU-time consumed to almost 115,000 cpu-hours (roughly 13 cpu-years).

5.2 Robustness

Each run of the HRA model is performed on an observation grid (with user-supplied spacing) and involves the sampling of probability distributions for the various HRA model parameters. Using finer grids and increased sampling will result in better estimation of the outcome (i.e. smaller confidence intervals), but also consumes more computational resources and consequently slows down the optimisation process. The execution time is roughly linear in the number of draws and quadratic in terms of the grid-size. It is, therefore, necessary to strike a balance between accuracy and overall computational time.

An additional issue is one of robustness, with respect to two levels of uncertainty –

- Stochastic uncertainty – given the stochastic nature of the HRA simulation, how big should the reduction in the various metrics be for them to be larger than the stochastic variability of the results? For a given choice of metric and initial condition, this applies to each of the 24 SPMI runs.
- Epistemic uncertainty – since it is not known which of the 24 logic-tree branches reflects the field reality, how do the gains due to optimisation compare to the HRA model's epistemic uncertainty?

The approach taken to address the stochastic uncertainty was as follows – for given grid density and number of draws ('catalogue count'), to first run the HRA model multiple times (21), with different seed value for random-number generator each time. The resulting set of values was then used as an input for a bootstrap-like estimation of the size of the 95% confidence interval relative to the estimated value. This

[9] A typical computing node in these data-centres has 24 Intel Haswell Xeon cores and 128GB onboard memory.

process was repeated, for each metric, for all the branches in the logic tree and then the mean value (for the relative width) was considered.

For example, Figure 5-5, which shows the results of such an analysis carried on a 500m grid, indicates that, for 40,000 catalogues, a reduction of maximal PGA by more than 5% could be considered robust[10]

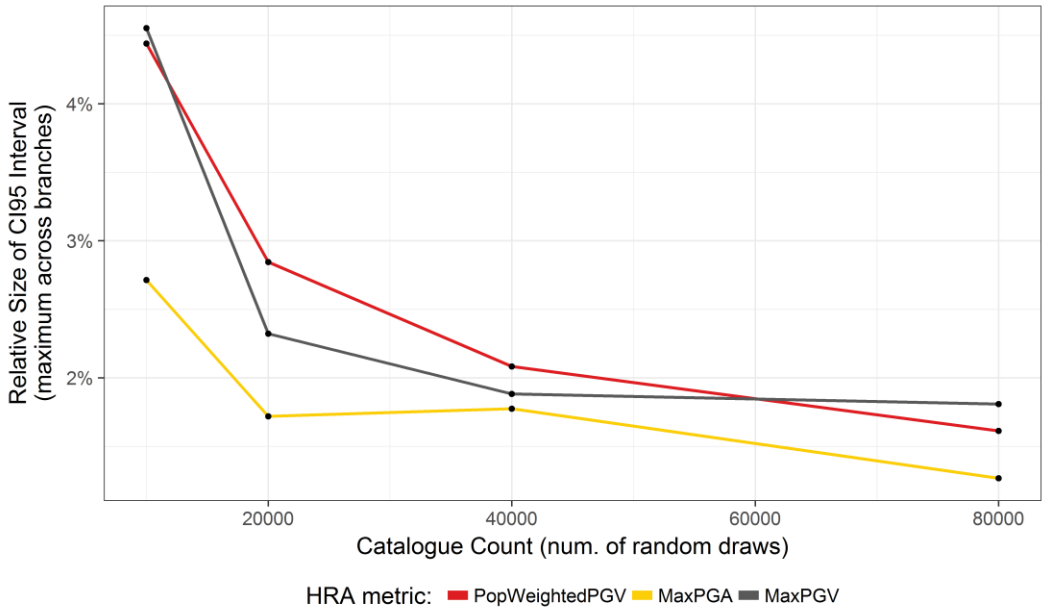


Figure 5-5: Sensitivity analysis to assess model stability

Qualitative assessment of the robustness with respect to epistemic uncertainty is somewhat more complicated, and was done by running the HRA simulation for the full logic-tree and comparing it to a reference case. The comparison considered the PGA/PGV difference map as well as the uniform hazard spectra in specific locations (e.g., Groningen, Loppersum, Ten Boer, etc.). In Figure 5-6, which shows an example of such a comparison, it can be seen that, although the epistemic uncertainty intervals do still overlap, there is a significant reduction of the hazard in Groningen and in Ten Boer.

A somewhat more rigorous approach to the assurance of robustness with respect to the epistemic uncertainty is to calculate, for each of the optimisation results, the expected value of optimisation metric. This is done by running the full logic-tree on the optimal distributions obtained from the 24 SPMI runs (the so-called 'assessment' mentioned in 5.1.4), and weighting the resulting metric values by the probabilities associated with each branch. The optimal distribution with the lowest expected value for the optimisation metric used is then selected (see Table 5-4 below).

[10] The main optimisation was done using a 1000m grid, with 50,000 catalogues.

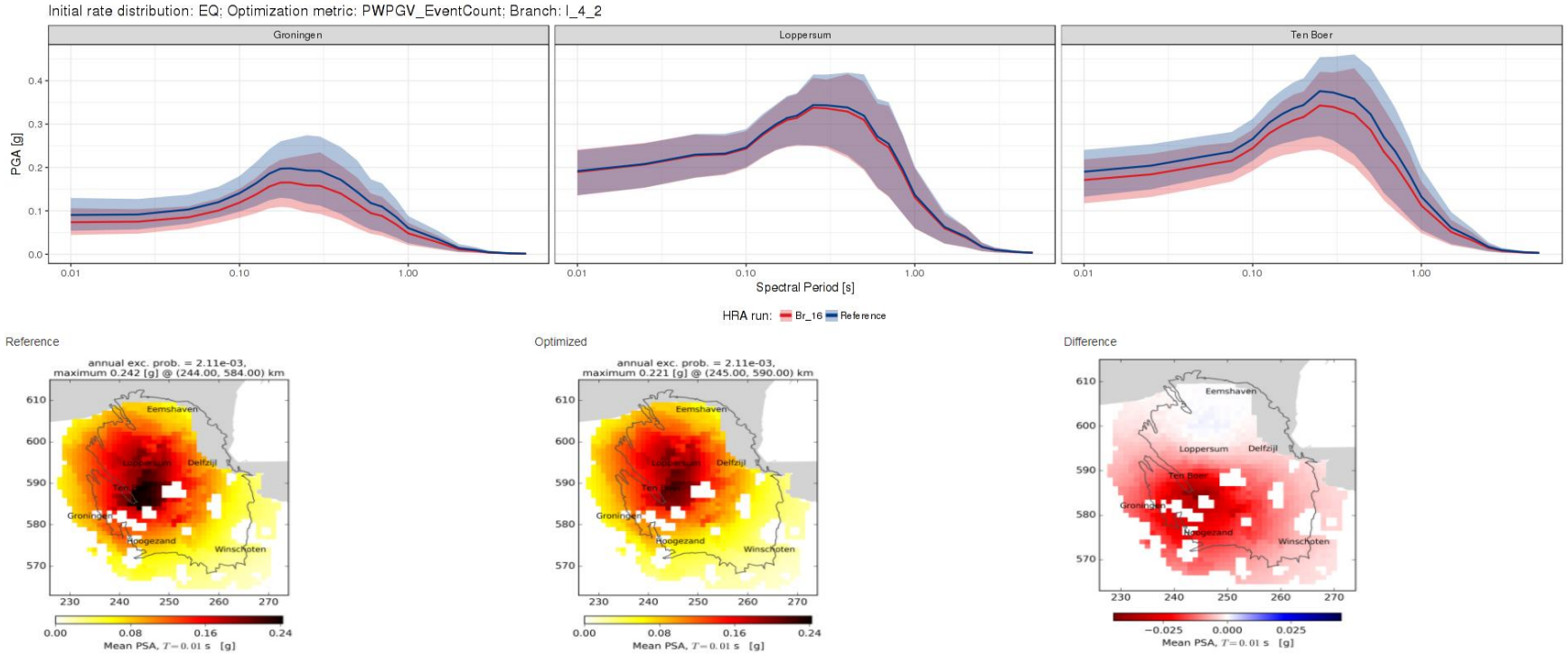


Figure 5-6 Sample HRA output visualization showing the PGA map and uniform hazard spectra (with the epistemic uncertainty) for selected locations. In the top panel, the reference and optimised solutions are plotted in blue and red, respectively (with shaded epistemic uncertainty intervals), and in the bottom, the PGA hazard (and difference) maps are plotted for the 475yr return period.

5.3 Initial areal trends in model results

The initial analysis of the results focused on establishing patterns in the thousands of simulation runs. Figure 5-7 gives an example, showing a scatter plot of maxPGV versus production fraction for each control, for branch 8 of the logic tree. Although directionally there are trends in the data, there is a big scatter and the trends do not always align between the different initial rates, which suggests that there is indeed a null space in the system. The epistemic uncertainty as captured by the full 24 branches of the Hazard tree further complicates the analysis. It was decided to embark on a more structured analysis of the model, Chapter 6.

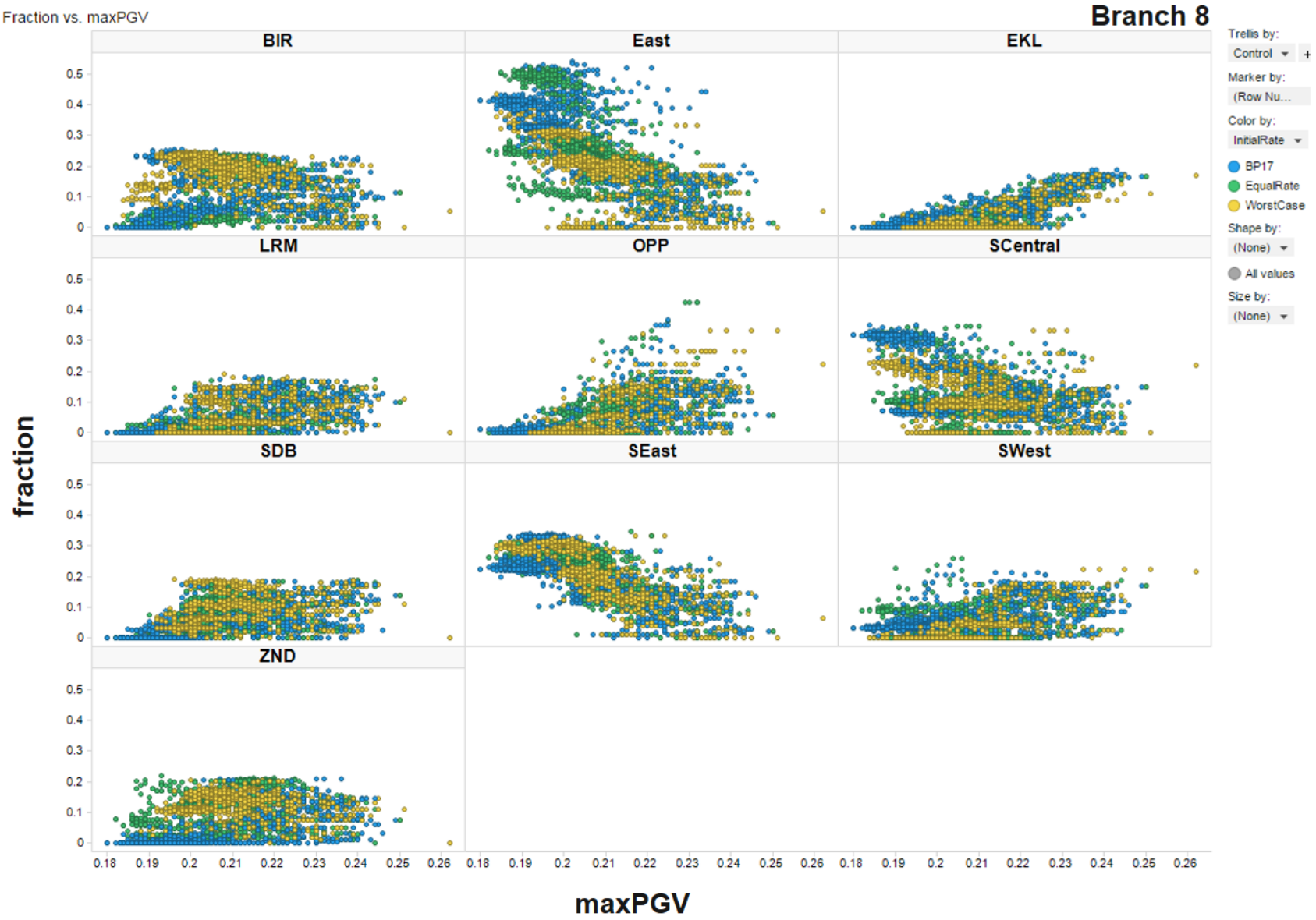


Figure 5-7: Scatter plot of maxPGV versus production fraction, by control, color coded by initial rate. Simulation results for Branch 8 of the logic tree

5.4 Optimisation results

A useful way to visualize the results is shown in Figure 5-10 (for optimisations using population weighted PGV as objective). Each figure contains the PGA difference maps (obtained by subtraction of the logic-tree means) between the reference case (BP17 production-distribution) and the 24 different optimisation outcomes (one per logic-tree branch). The full set of results for each branch for the different objectives, along with corresponding production splits and PGA difference maps, are given in **Error! Reference source not found.**. All the results reported in this section and in **Error! Reference source not found.** are based on using the BP17 production-distribution as starting point for the optimisation. While the optimised production splits do vary somewhat depending on the starting point, the overall reduction in seismicity metrics are in good agreement.

Comparison of the individual branch results for a given optimisation metric (i.e., the 24 results within Figure 5-10) shows that, for the most part, the choice of the driving branch does not have significant effect on the outcome. On the other hand, the choice of optimisation metric can, and does, result in noticeable changes to the hazard map.

As explained above, the selection of the optimal solution from amongst the optima obtained by the SPMI runs ("branch-optimal" solutions) is based on the expected value for the optimisation metric used. Table 5-4 below summarizes these results, Figure 5-8 shows the corresponding production distributions, and Figure 5-11 to Figure 5-12 show the PGV and PGA difference maps with respect to the reference case.

As mentioned in Section 1.5.2, all regions of the Groningen field are in pressure communication. Re-distributing production between the regions will change the areal distribution of pressure decline, with subsequent impact on compaction and seismicity. However, all possible production configurations will still in sum deplete the entire field. In Figure 5-9 the drainage areas of three of the optimisations are compared – maximum PGV, population weighted PGV and hybrid optimisation. The different optimisations result in distinctively different drainage patterns, but in all cases the entire field is covered.

In Figure 5-11 and Figure 5-12 the metrics that are 'global' (i.e. ones which do not include any spatial weighting) result in an overall reduction of the hazard (the PGA and PGV maps show reduction throughout the area). On the other hand, the pwPGV-driven optimisation yields a greater reduction in the populated areas due to the spatial weighting, but that is offset by a slight increase in PGA/PGV values in the less-populated areas. Finally, the hybrid approach, which combines a 'global' metric (event count) with the population-weighted PGV.

The spatial impact of the optimisation is visualized in Figure 5-13, which shows the optimum production fractions as compared to their initial distribution. The results roughly align with the outcome from the Mores stand-alone optimisation (section 4.5) but do show a spread depending on the initial distribution, further indicating the presence of a null space. Figure 5-14 highlights the spread in the optimum production fractions as a function of the objective function, whereas Figure 5-15 shows the spread as a function of the initial rate. Note that maxPGA and maxPGV both represent a hazard metric and yield similar optimisation results, hence maxPGA was omitted from Figure 5-13 to Figure 5-15.

Table 5-4: Summary of optimisation results

	HRA Metric Values				HRA Metric Values			
	pwPGV	Max PGA	Max PGV	Event count	pwPGV	Max PGA	Max PGV	Event count
Ref	**0.052**	**0.246**	**0.143**	**115**				
Event Count	0.045	0.225	0.133	**95**	-13%	-9%	-7%	**-17%**
Max PGA	0.045	**0.226**	0.135	96	-13%	**-8%**	-6%	-16%
Max PGV	0.047	0.228	**0.132**	99	-9%	-7%	**-8%**	-14%
pwPGV	**0.044**	0.230	0.143	99	**-15%**	-6%	0%	-14%
Hybrid	**0.043**	0.228	0.138	**95**	**-16%**	-7%	-4%	**-17%**

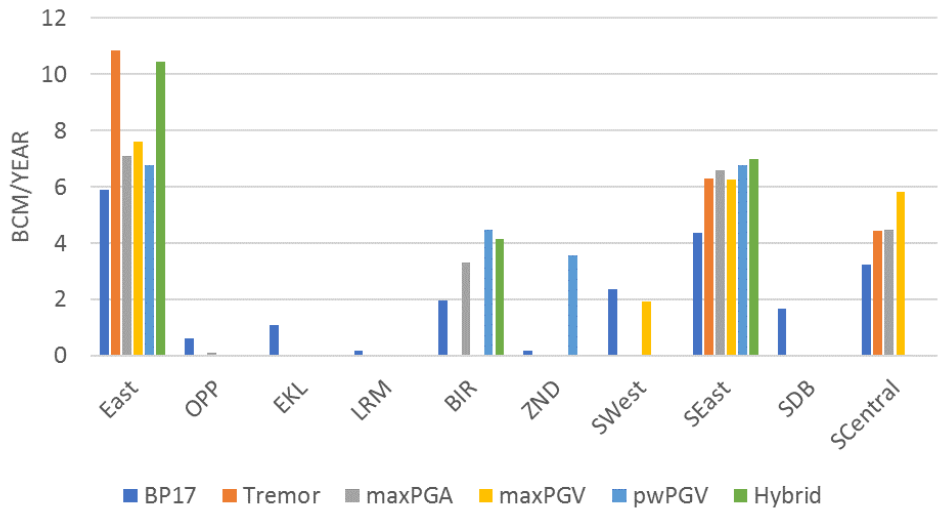


Figure 5-8: Allocated production fraction for each optimisation metric with reference to BP17, by control.

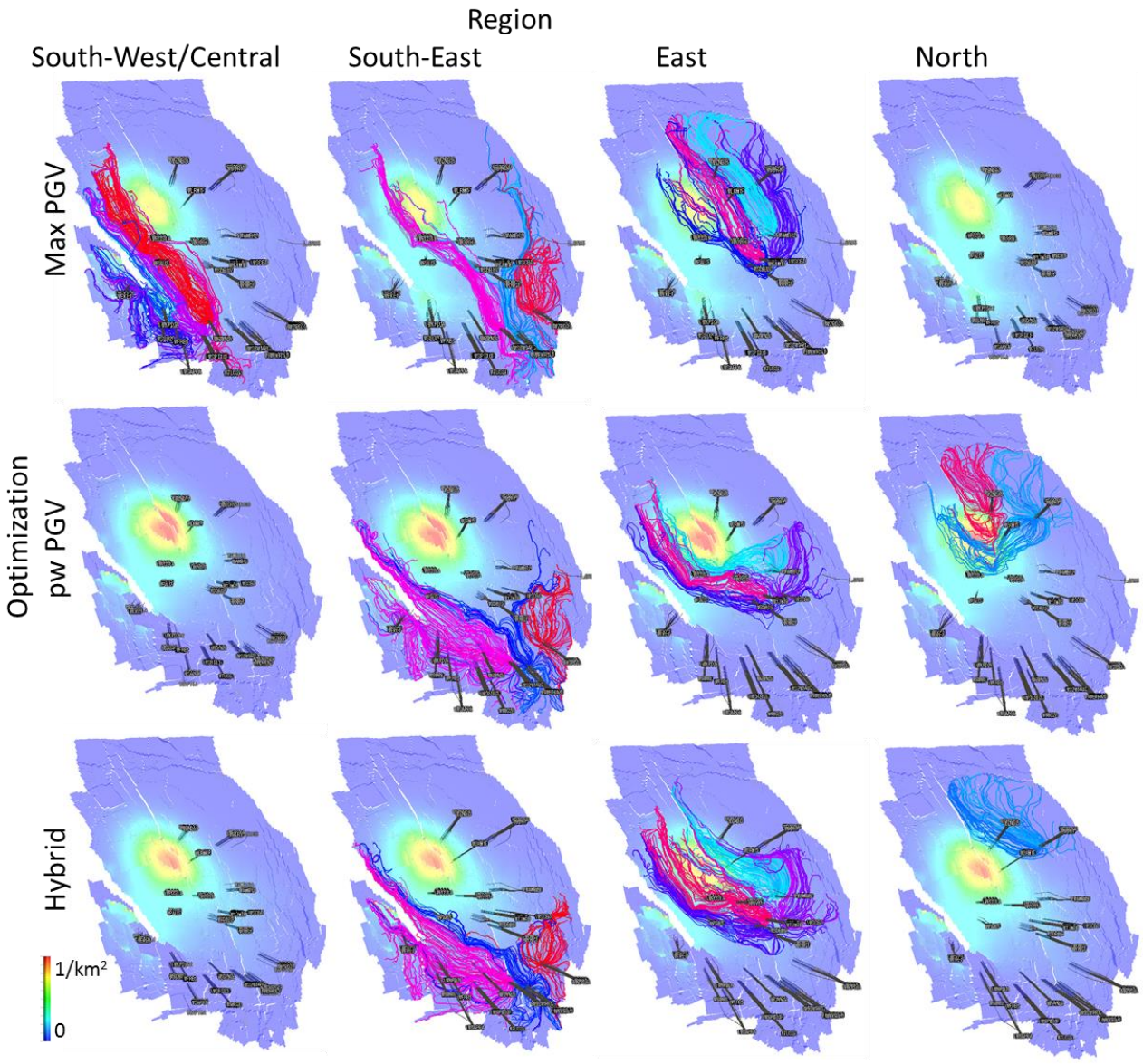


Figure 5-9: Distribution of streamlines resulting from max PGV, population weighted PGV and hybrid optimisation. Streamlines are colored by arriving producer, constrained to production regions (optimisation controls) – South West/Central, South-East, East and North. Please refer to Figure 1-7 for a definition of the regions. Streamlines are based on production distribution at the end of the optimisation period, 31-12-2022. Underlying property is cumulative modelled earthquake density over the optimisation period, 2018-22. The Mores stand-alone tremor model is used for the density plots.

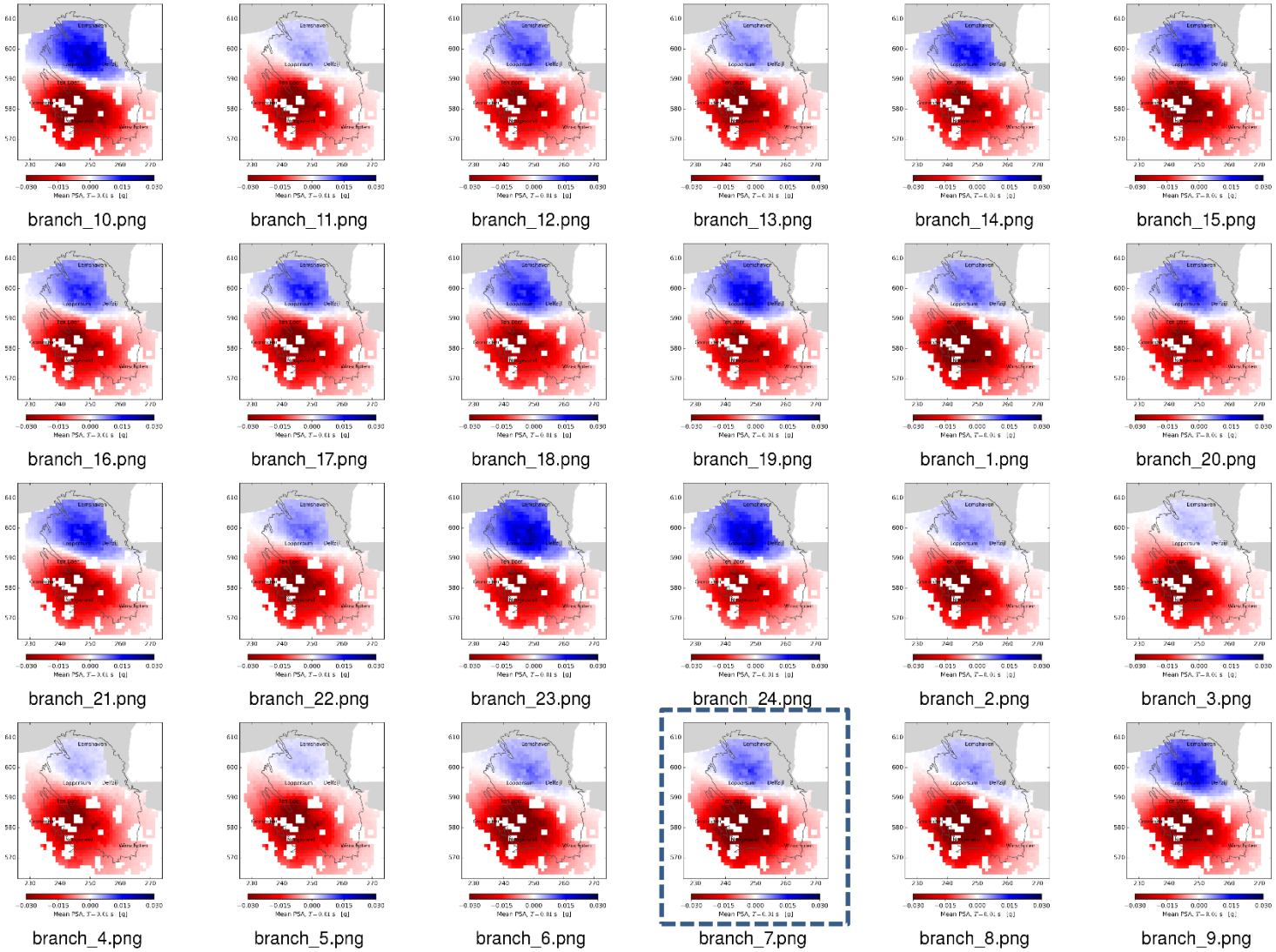
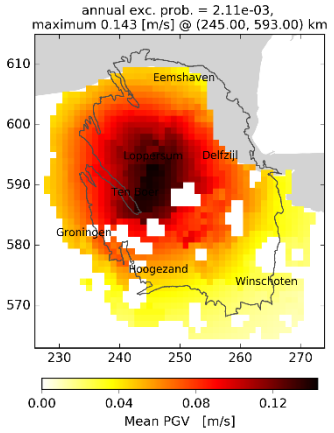
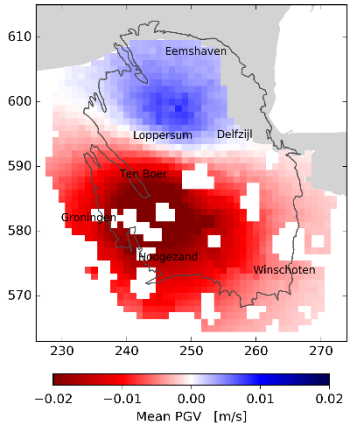


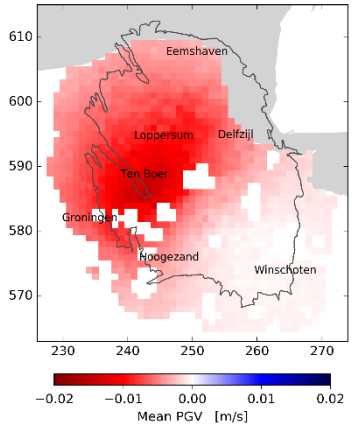
Figure 5-10: Mean PGA hazard difference maps (scaled [-0.03,0.03]) for Population-weighted PGV driven optimisation, at average 0.2% annual chance of exceedance (1 in 475 years) from 1/1/2018 to 31/12/2022 (optimal solution indicated)



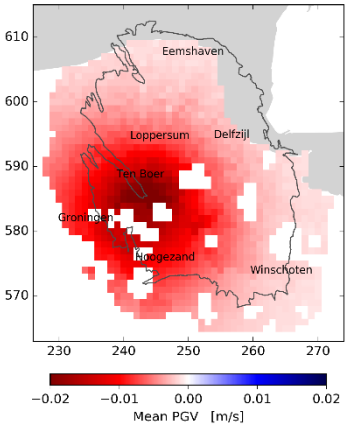
Reference (Max PGV: 0.143 m/s)



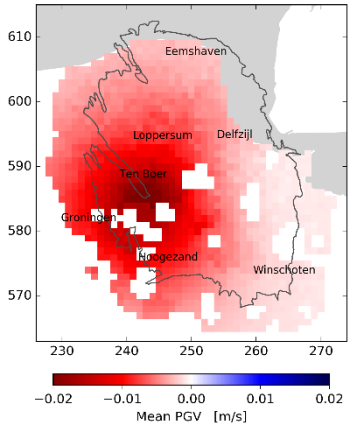
Population-weighted PGV (Max PGV: 0.146 m/s)
Range: (-0.031,0.001)



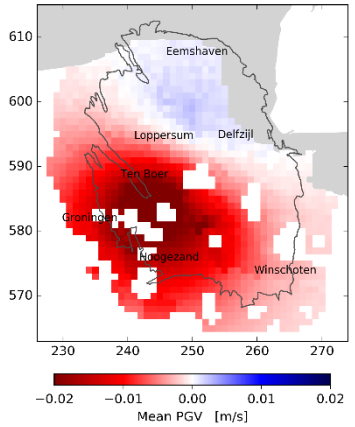
Maximal PGV (Max PGV: 0.132 m/s)
Range (-0.016,0.0)



Maximal PGA (Max PGV: 0.135 m/s)
Range: (-0.042,0.0)

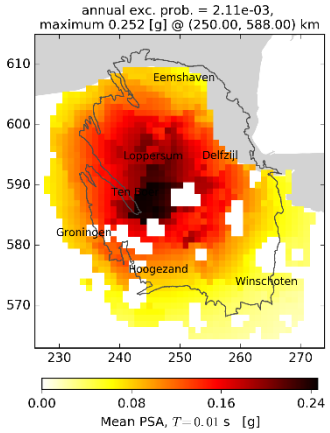


Event-count (Max PGV: 0.133 m/s)
Range: (-0.0244,0.0)

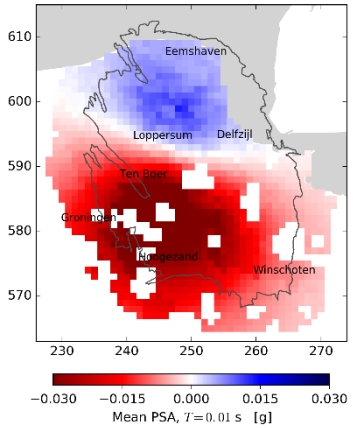


Hybrid(pwPGV+EventCount) (Max PGV: 0.138 m/s)
Range: (-0.0223,0.0001)

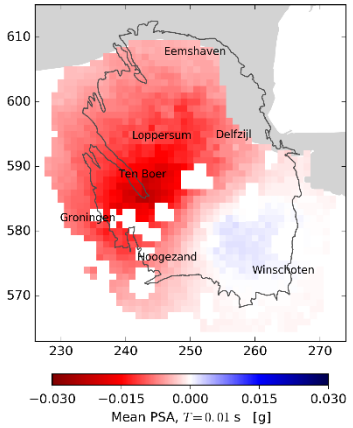
Figure 5-11: PGV difference maps for the optimal solutions for each metric



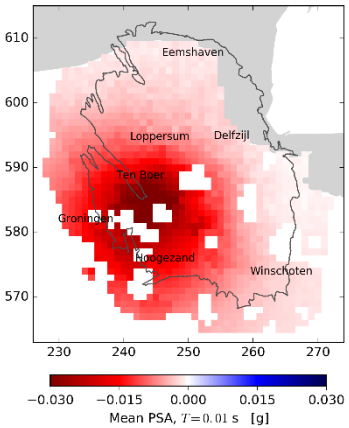
Reference (Max PSA: 0.252g)



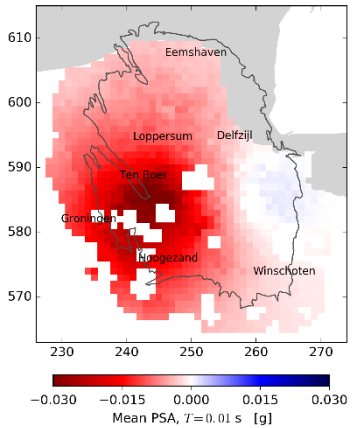
Population-weighted PGV (Max PSA: 0.236g)
Range: (-0.055,0.004)



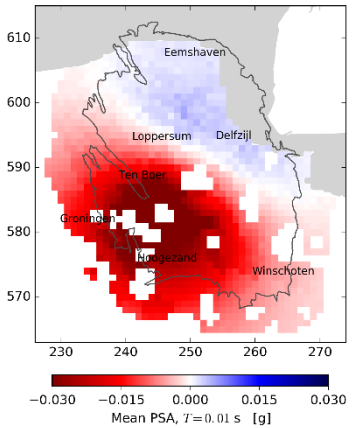
Maximal PGV (Max PSA: 0.238g)
Range: (-0.026,0.003)



Maximal PGA (Max PSA: 0.232g)
Range: (-0.0426,0.0)



Event-count (Max PSA: 0.234g)
Range: (-0.041,0.002)



Hybrid(pwPGV+EventCount) (Max PSA: 0.232g)
Range: (-0.038,0.001)

Figure 5-12 PGA difference maps for the optimal solutions for each metric

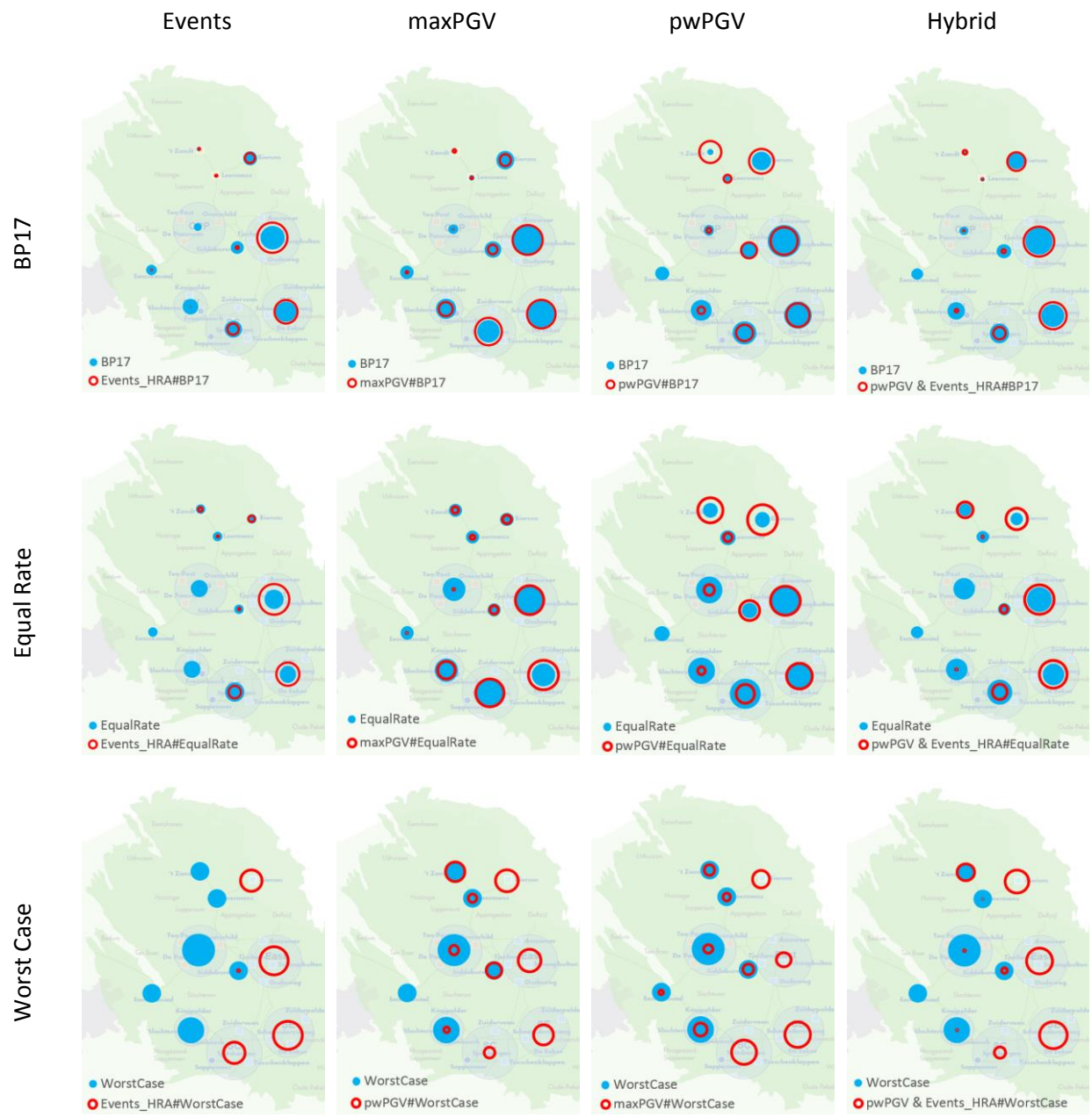


Figure 5-13: Spatial display of optimisation results by objective function (columns), per initial production distribution (rows). The production fractions are indicated as a bubble map, with the size of the bubble scaled to the production fractions. In blue the initial distribution, and in red the optimisation results.

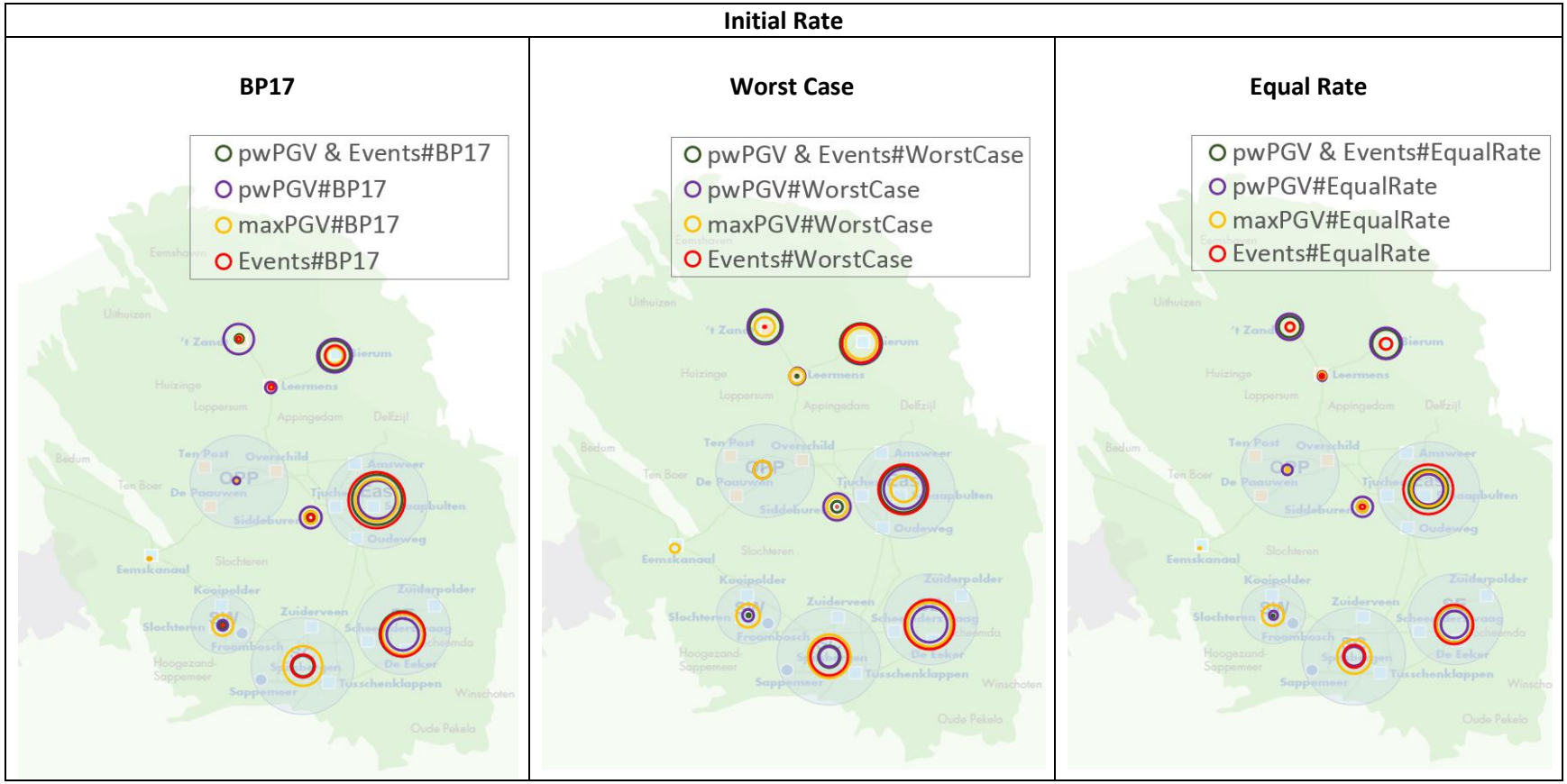


Figure 5-14: Bubble plots showing the optimised production distribution for each Objective Function, trellis by initial rate scenario (BP17, WorstCase, EqualRate).

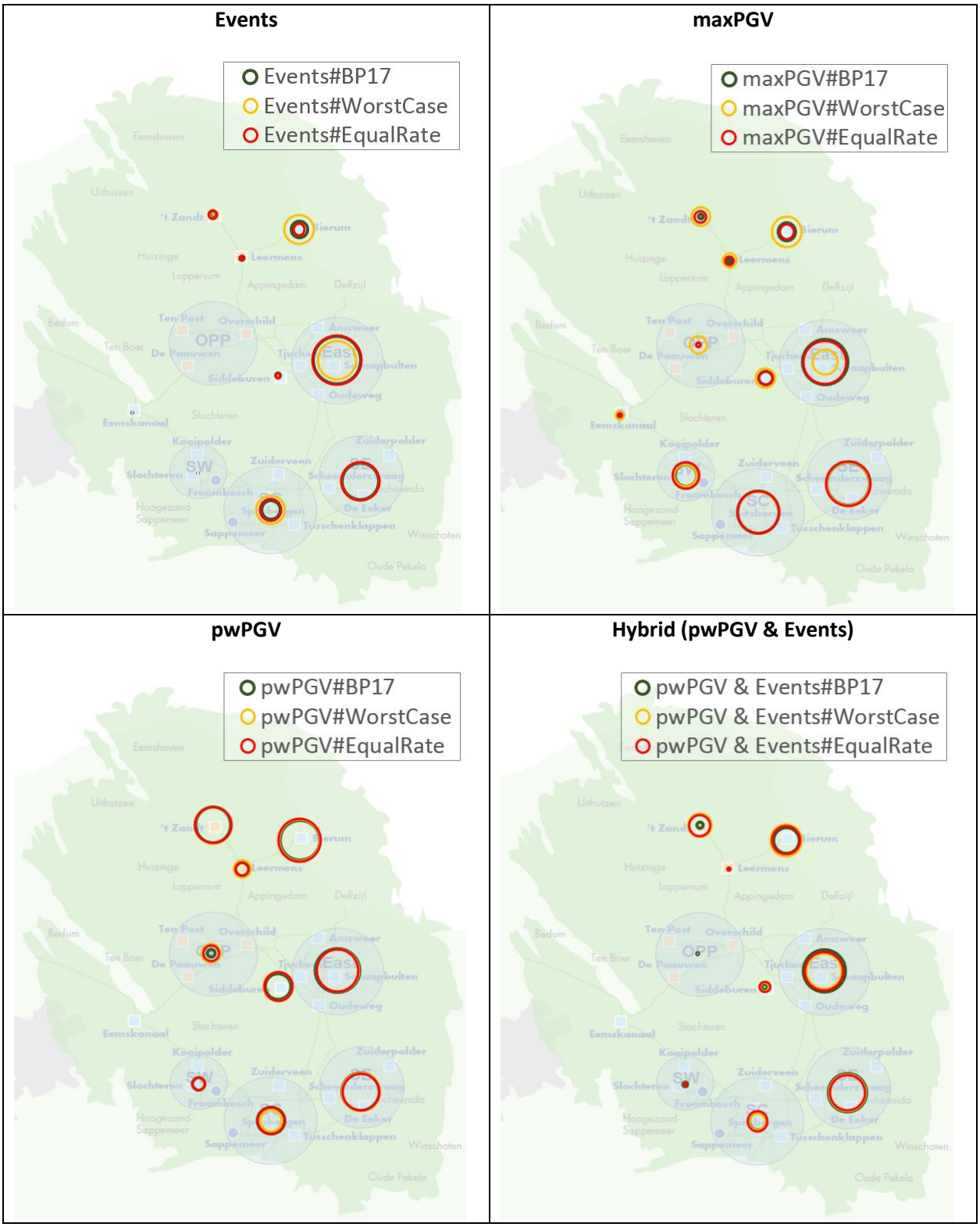


Figure 5-15: Bubble plots showing the optimised production distribution for each initial rate scenario (BP17, WorstCase, EqualRate), trellis by Objective Function.

6 Data driven analysis of model response

To enhance and complement the understanding of the optimisation results, a structured analysis of the model response was made.

6.1 Exploring the Model Response Across Different Branches

In order to be able to explore and reasonably characterize the model response with respect to changes in the inputs of a complex model that cannot be easily analytically described, one approach is to approximate the complex model with a simpler but still reasonably accurate model that can be more directly analysed. To create this "proxy model", it is necessary to create a sufficient number of control and output pairs to which the proxy model is fit. To be able to make statements about the overall quality of the model, it is necessary that the model is based on a representative set of control output pairs that capture enough of the dynamic behaviour of the original model.
To achieve this, the control space was uniformly randomly sampled and it was tested whether the out of sample prediction performance of the proxy model exceeded a certain threshold that is deemed sufficient for the modelling purposes. This threshold is chosen to be close to the irreducible uncertainty level that is due to the probabilistic nature of the HRA. Once this threshold is exceeded, no additional samples are generated. One proxy model is created for each of the HRA branches. A conceptual overview of the workflow for one individual branch of the HRA is shown in Figure 6-1.

After briefly introducing the mathematical notation used throughout this subsection and associated appendix, it is explained how uniform samples can be generated in the control space. Next, the type of proxy model used is explained, and how its performance is tested. Eventually, it is explained how variable importance and partial dependence plots based on the model are created to capture the model response. Finally, the results are given as obtained for the different HRA branches.

A more general introduction on the topic, the employed modelling techniques as well as the associated diagnostic tools can be found in Appendix F.

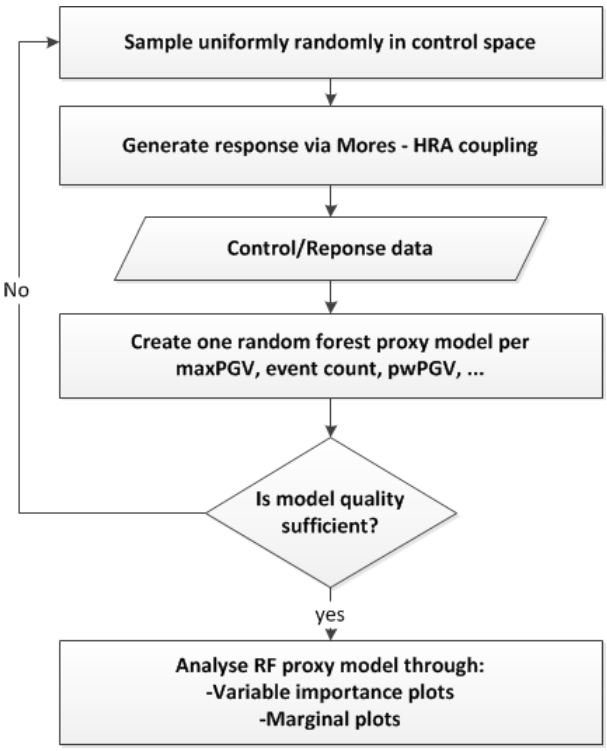


Figure 6-1: Sketch of proxy workflow for one HRA branch

6.2 Mathematical Notation and Definitions

The mathematical notations and abbreviations used are as follows:

m	number of HRA branches ($m > 0$)
n	number of controls, i.e. the number of cluster groups that contribute to the total production ($n > 0$)
p_j	number of samples for branch j, with $1 \leq j \leq m$ and $p_j > 0$
c_{ij}	fractional contribution of cluster i for branch j to the total normalized production of 1, $(0 \leq c_{ij} \leq 1)$

Furthermore, for $1 \leq j \leq m$ the equality $\sum_{i=1}^{n} c_{ij} = 1$ holds. For the sake of simplicity, in the following, the focus is on one individual branch at a time, and the corresponding branch index is omitted in the notation.

6.3 Uniformly Sampling from a Convex Polyhedron

First, it is noted that the admissible control strategies, due to the constraint of $\sum_{i=1}^{n} c_i = 1$, all lie on a unit $(n-1)$-simplex. A simplex is the generalization of the concept of triangles and tetrahedra to higher dimensions. See Figure 6-2 for an illustration of several admissible combinations with 3 controls which lie on a triangle. The theory for sampling uniformly from a unit simplex is well established and for instance described in [17] p. 568. A rather simple algorithm for drawing uniformly random samples that is also intuitive works as follows:

1. Draw $n-1$ samples uniformly from the interval $[0,1]$. Call the resulting set S.

2. Add the numbers 0 and 1 to the set S.
3. Order the set S in increasing order and denote its ordered elements by s_i with $1 \leq i \leq n+1$.
4. Let $c_i = s_{i+1} - s_i$ for $1 \leq i \leq n$.

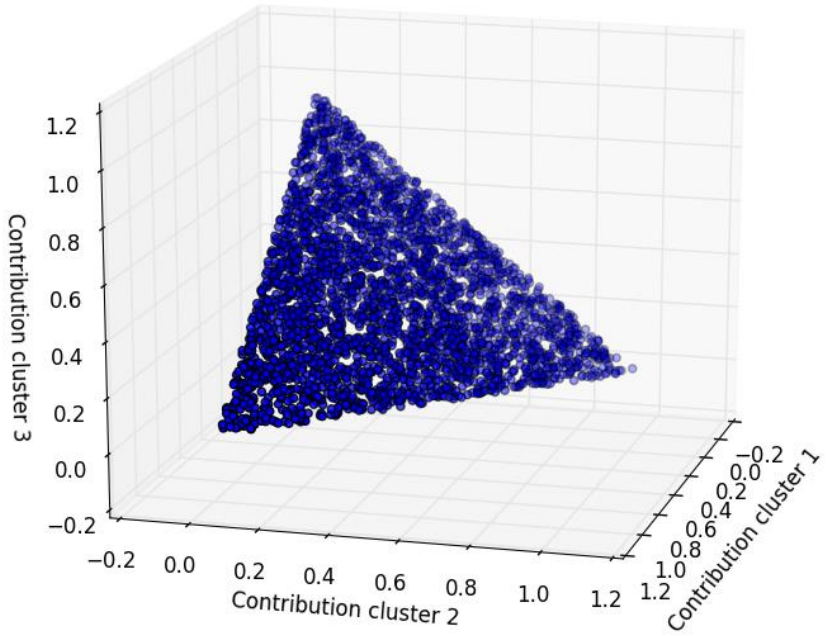


Figure 6-2: Admissible combinations for 3 controls lie on 2-simplex

One can show that the algorithm described above will create uniformly random samples in the unit $(n-1)$-simplex, see [17] p. 568 Theorem 2.1. The computational cost to create p_j samples for a particular branch j is in $O(p_j \log p_j)$, which is negligible compared to the time it takes to run the associated Mores model in combination with the HRA. A set of 3000 samples that have been generated using this approach is contained in Figure 6-2.

Remark:
The strategy of drawing n times uniformly from the interval $[0,1]$ and then normalizing by the sum to obtain c_i such that $\sum_{i=1}^{n} c_i = 1$ and repeating this procedure p_j times will yield samples that are not uniformly distributed in the control space. See Figure 6-3 for an illustration. Extreme combinations of parameters are sampled with a lower probability which may mean that the control space is not properly covered. This may be even more of an issue in a high dimensional space, as with $n = 10$, and/or if the model response is complex towards the extreme points of the control space.

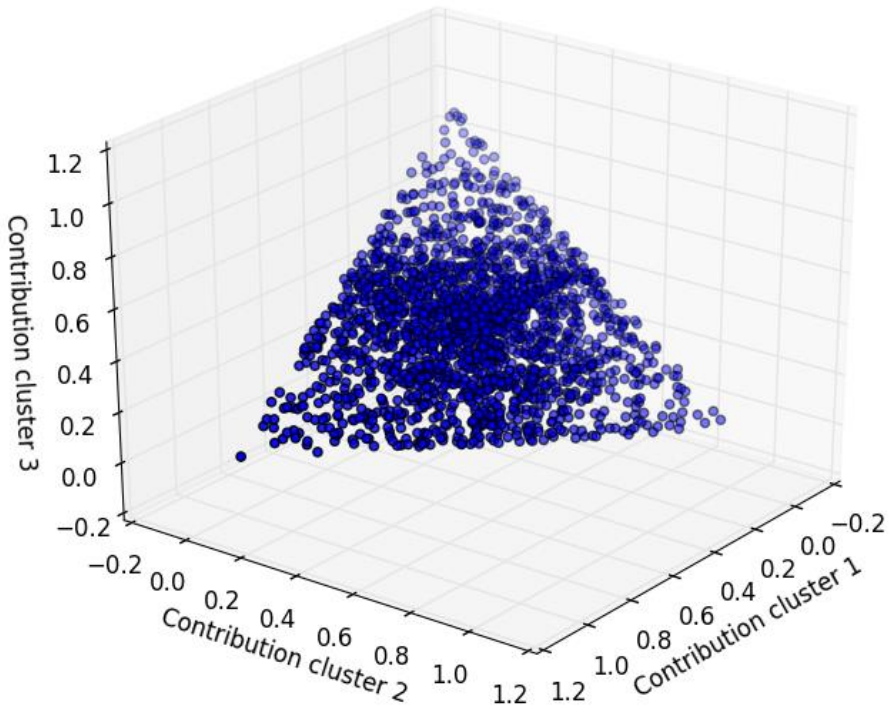


Figure 6-3: Non-uniform sampling of a 2-simplex

The actual situation at hand is slightly more involved than just sampling from a unit simplex, since some of the controls themselves are subject to boundary conditions, which are based on the maximal contribution that individual cluster can make over the forecasting horizon for which the contribution of the cluster is kept constant. These upper boundaries have been determined empirically based on a set of several prediction runs. The lower boundaries are in place to ensure the minimal operational constraints of the individual clusters are met.

The following boundary conditions were applied:

$$0.01 \leq c_1 \leq 0.56 \qquad 0.01 \leq c_2 \leq 0.54 \qquad 0.01 \leq c_3 \leq 0.20 \qquad 0.01 \leq c_4 \leq 0.19 \qquad 0.01 \leq c_5 \leq 0.24$$

$$0.01 \leq c_6 \leq 0.20 \qquad 0.01 \leq c_7 \leq 0.30 \qquad 0.01 \leq c_8 \leq 0.41 \qquad 0.01 \leq c_9 \leq 0.18 \qquad 0.01 \leq c_{10} \leq 0.44$$

The resulting solution space that one obtains if those boundary conditions are in place is a convex polyhedron. For illustrative purposes consider the case with 3 controls which have the following boundary conditions:

$$0.2 \leq c_1 \leq 0.60 \qquad 0.2 \leq c_2 \leq 0.60 \qquad 0.10 \leq c_3 \leq 0.90$$

A visualization of the solution space is contained in Figure 6-4.

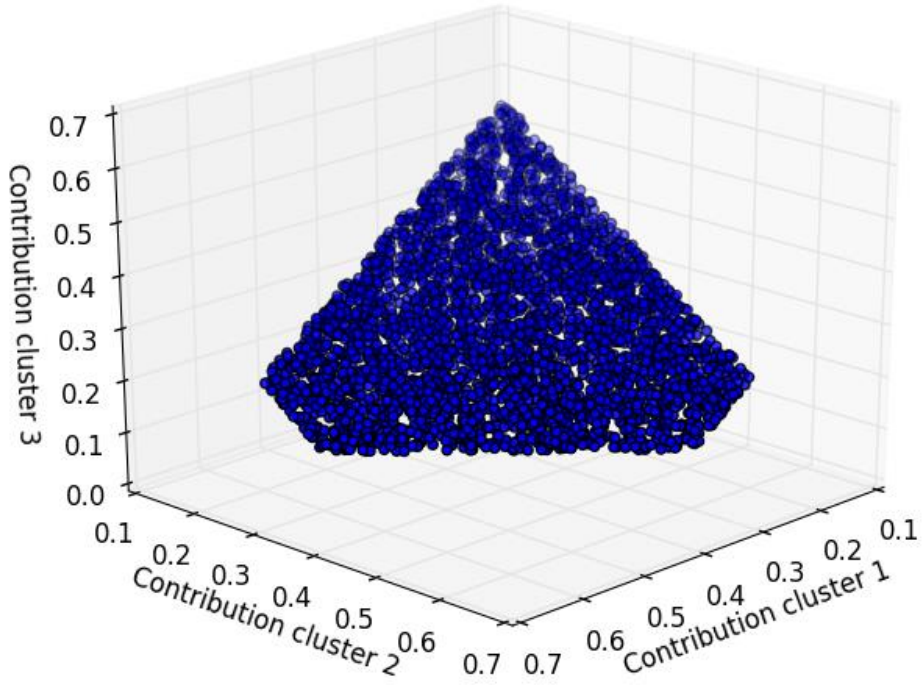


Figure 6-4: Solution space for 3 controls with constraints is a convex polyhedron

The strategy followed to sample uniformly from a convex polyhedron is to sample uniformly from the associated unit simplex and to discard all samples which do not satisfy the boundary conditions on the c_i as specified above. This rejection sampling strategy is straightforward to implement. Even though more direct techniques would exist that are computationally less costly, the cost to generate a few 1000 samples for 10 controls is still negligible (in the order of seconds) compared to the cost of running the coupled Mores and HRA models. The sampling depicted in Figure 6-4 was obtained using this rejection sampling strategy.

6.4 Building a Suitable Proxy Model

Having established how the sampling procedure works, it is now explained how to determine the number of samples needed (per branch) to adequately capture the behaviour of the Mores-HRA coupling through a proxy model. The chosen class of proxy models are Random Forests (RFs). Other models may have worked equally well; however RFs have the following features that contribute to their particular utility in this case:

- Can capture non-linear effects and interactions between controls
- Good out-of-box performance without tuning of algorithm parameters
- Computational efficiency
- Diagnostic tools that allow to extract information like variable importance and partial dependence of model response on a subset of controls.

A more detailed explanation about how these models are created with some illustrative examples can be found in Appendix F.3. Additional theoretical information can be found, for instance, in Reference [18].

As part of the iterative procedure that is sketched in Figure 6-1, a RF model is trained on those available control response pairs that manage to meet a minimally aspired production threshold of 21.4 Bcm/year. Even though constraints were already applied on the c_i to ensure that (most of the) combinations would achieve the minimal production threshold of 21.4 Bcm/year, some of the more extreme combinations may still drop out.

The predictive quality of the model is evaluated out-of-sample. Further details on the methodology can be found in Appendix 0. Once the model quality reaches a certain threshold value which comes close to the uncertainty that is inherent to the HRA (between 1% and 2% depending on the HRA branch), the number of samples is considered sufficient. Through this procedure, it was established that around 3,000 uniform samples per branch are sufficient to build a suitable proxy model. The exact number of control configurations that managed to make the minimal production constraints and the out-of-sample performance of the respective random forest models for pwPGV that were trained on the available samples is shown in Table 6-1. The estimated standard error for the reported out of sample R^2 was between 0.01 and 0.02, hence there is little variability in this measure of model performance.

Table 6-1: Number of control configurations that lead to at least 21.4 Bcm/year and the associated R^2 and Normalized RMSE of the RF models for pwPGV. The normalized RMSE is in the range of uncertainty inherent to the HRA of around 1-2%.

Branch number	1	2	3	4	5	6	7	8	9	10	11	12
Number of samples	2322	2354	2411	2362	2371	2365	2359	2349	2317	2335	2362	2364
Out of sample R^2	0.96	0.96	0.96	0.96	0.95	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Normalized RMSE in %	1.45	1.42	1.45	1.50	1.52	1.51	1.53	1.58	1.34	1.33	1.35	1.36

Branch number	13	14	15	16	17	18	19	20	21	22	23	24
Number of samples	2343	2337	2351	2337	2369	2326	2385	2307	2369	2395	2407	2342
Out of sample R^2	0.96	0.96	0.96	0.96	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
Normalized RMSE in %	1.36	1.35	1.34	1.34	2.01	2.10	2.18	2.17	2.33	2.27	2.33	2.21

6.5 Evaluating Model Response Through Variable Importance Plots

One common technique, that is used to estimate the overall contribution (including non-linear and interaction effects) of a variable on the model response of a random forest model, are so called variable importance plots. Some details about how they are computed can be found in Reference [18] and Appendix F.4. The permutation based estimate of mean increase in MSE measure is utilized, as described in more detail in Appendix 0. For such plots to be meaningful, it is important that the controls are uncorrelated. If that is not the case, the contribution of two controls on the model response will be commingled and hence an individual relationship between a control and model response can no longer be established. Since a uniformly random sampling setup is used, sufficient independence of controls may be assumed. This is also illustrated by Table 6-2 that contains the (Pearson) correlations between the controls in the data set for branch 1. On average, a slight negative correlation can be observed, which is due to the fact the controls need to add up to 1. Hence an increase in one control will on average need to lead to a slight decrease in other controls. Since the correlations between the controls are only based on the random sampling setup, very similar observations hold for the other branches.

Table 6-2: Correlations between controls for samples from branch 1

Branch 1 correlations										
	BIR	**EKL**	**East**	**LRM**	**OPP**	**SCentral**	**SDB**	**SEast**	**SWest**	**ZND**
BIR	1.00	-0.06	-0.16	-0.05	-0.14	-0.10	-0.06	-0.10	-0.05	-0.06
EKL		1.00	-0.14	-0.03	-0.05	-0.11	-0.05	-0.06	-0.06	-0.03
East			1.00	-0.12	-0.25	-0.18	-0.13	-0.21	-0.16	-0.14
LRM				1.00	-0.13	-0.07	-0.02	-0.06	-0.06	-0.06
OPP					1.00	-0.20	-0.13	-0.21	-0.15	-0.11
SCentral						1.00	-0.05	-0.20	-0.16	-0.08
SDB							1.00	-0.09	-0.02	-0.02
SEast								1.00	-0.08	-0.06
SWest									1.00	-0.07
ZND										1.00

An example plot is contained in Figure 6-5, which shows a variable importance plot for a RF model of branch 1 that predicts pwPGV. The variables, are ordered decreasingly with respect to their cumulative impact on pwPGV. The associated standard error in the estimates are indicated by the whiskers on the bars. Note that this type of plot does not allow to draw any conclusions about the nature of the relationship between a variable and the objective function, as in "an increase/decrease in production from region X leads to an increase/decrease in metric Y". However, this plot gives an indication about the order in which regions should be investigated based on their impact on the objective function. How the average effect of changes in a control can be assessed will be discussed in section 6.6.

Additionally, variables that are deemed insignificant in this representation, i.e. variables whose estimated increase in MSE is close to the standard error of the measurement are variables that can be changed without significant changes in the general model response. Based on Figure 6-5 this would imply that changes to controls ZND, LRM, and SDB have virtually no effect on the model response for branch 1. Since the model is known to be a good proxy for the coupled Mores HRA setup, it can be concluded that those parameters are essentially in the model null-space.

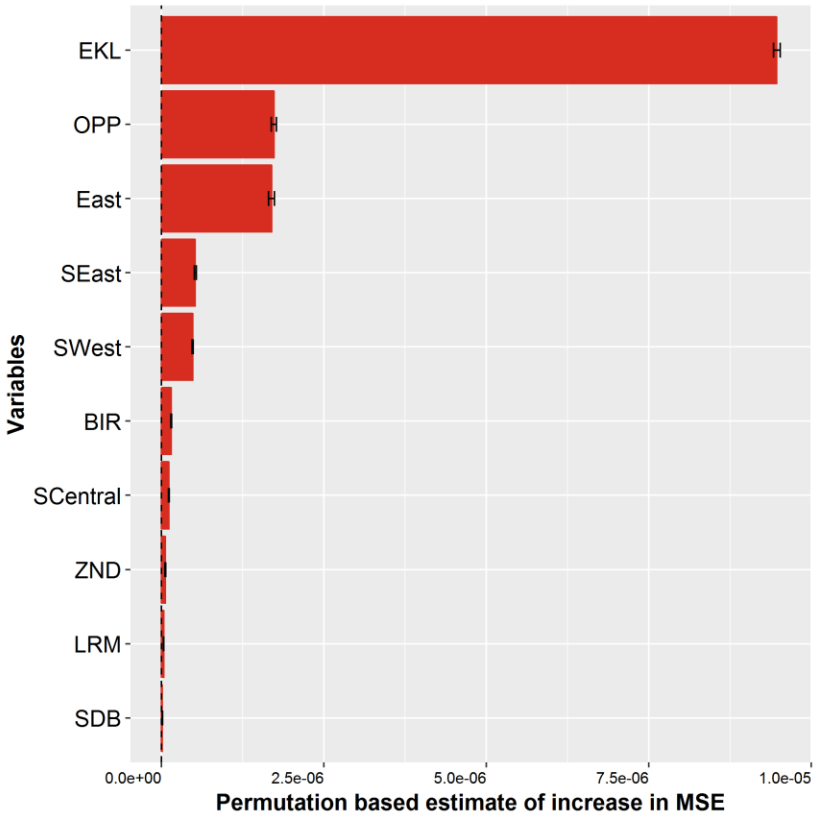


Figure 6-5: Variable importance estimates using the permutation based increase in MSE of pwPGV with associated standard errors for branch 1. The variables are ordering decreasingly with respect to their cumulative impact on pwPGV.

6.6 Evaluating Model Response Through Partial Dependence Plots

In order to be able to visualize and grasp the effects of individual groups of variables on the model response, so-called partial dependence plots can be used. Again, further details can be found in Appendix F.5 and in Reference [18]. As mentioned before, for those plots to be meaningful it is important to ensure that the controls are not highly correlated. A set of example plots for the average model response on changes in individual controls is shown in Figure 6-6.

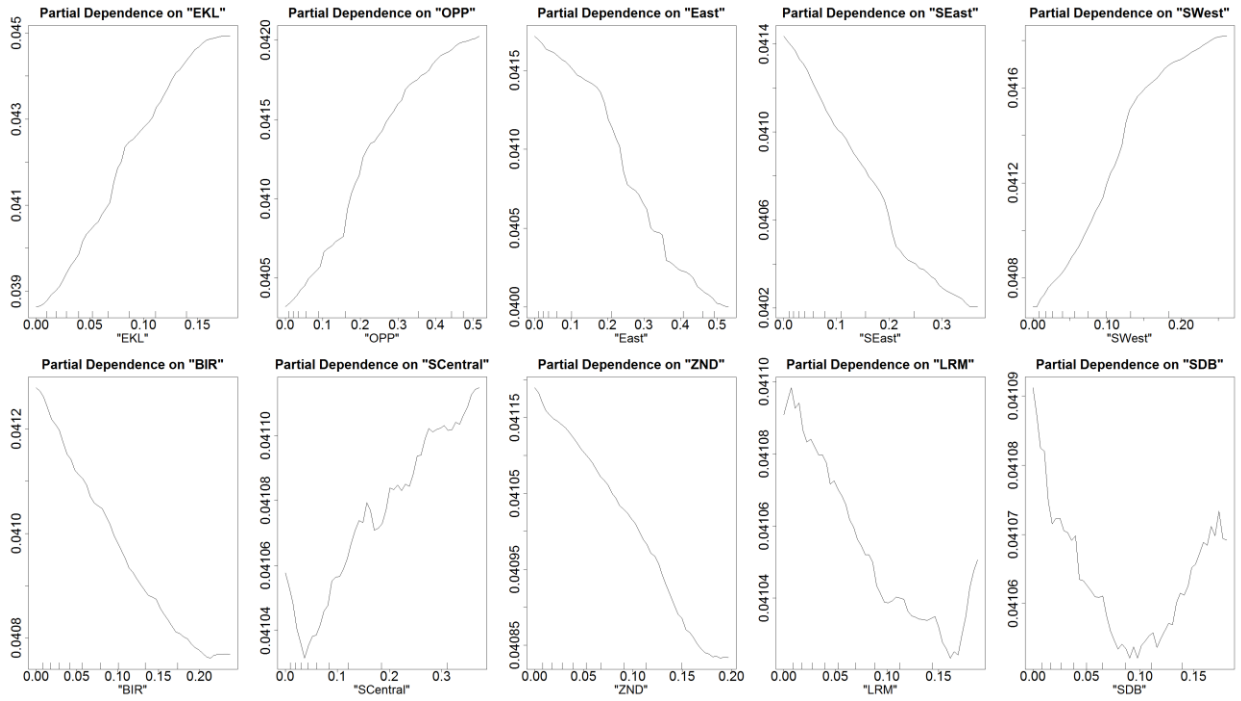


Figure 6-6: Partial dependence plots for average model response of pwPGV with respect to changes in individual controls for branch 1. Plots are ordered by the variable importance as indicated in Figure 6-5. Note that the y-scale for each plot is different.

6.7 Partial dependence plots for all branches

For any (deterministic) model realisation, the total of the production fractions over all controls adds up to 1. Changing the production fraction for a single control would involve having to make a choice on how to offset that change by an opposite change in the total production fraction from all other controls. However, any specific choice would impact on the outcome. The power of the partial dependence plots is that it allows to evaluate the impact of a single control with respect to an "averaged" response of the rest of the system.

A positive slope means that an increase in the production fraction of that specific control on average increases the value of the objective function (i.e. increases the Hazard/Risk). A negative slope means that increasing the production fraction on average leads to a reduction in the value of the objective function (i.e. getting closer to the optimum). A zero slope means that the value of the objective is essentially not affected by the production fraction of that control.

Figure 6-7 gives a trellis of the partial dependence plots for all 10 controls with respect to pwPGV. For each control, all 24 branches of the Hazard tree are given, colour coded from branch 1 in blue to branch 24 in red. It can be observed from Figure 6-7 that the slopes of these partial dependence curves are highly consistent across all 24 branches: they are either all positive, or all negative, are all zero. All slopes are also roughly linear, hence the relationships can be approximated by fitting a linear trend line through each slope.

Figure 6-8 plots the corresponding best fit linear slopes for all controls, again colour coded by branch. Although there is a bit of a spread for the more influential controls (i.e. the controls with the larger slopes, either positive or negative), the consistency of the dependencies clearly comes out: either all positive, negative, or zero. This provides a clear guidance for the optimisation. There is an epistemic uncertainty range, which is span by the branches of the hazard tree. Despite the fact that we don't know which branches best approximate the truth, we can still optimise because all branches reflect the same directionally. Average values for the slopes of each control across the 24 branches were included in Figure 6-8.

The partial dependence analysis was done for all five objectives (each for all 24 branches), showing highly similar (linear-slope) behaviour. In Figure 6-9 the averaged slopes for all controls are superimposed on a map in a traffic light style. The red part of the colour spectrum reflects the positive slopes. For those controls, an increase in the production fraction will increase the objective function (i.e. increase the hazard or risk). The green part of the colour spectrum reflects the negative slopes, hence increased production from those controls will reduce the hazard or risk. The white controls are indifferent, more or less production will not impact the hazard or risk.

Some clear patterns stand out from Figure 6-9 across the various objective functions. All agree that the Eemskanaal production cluster has a profound negative impact on the seismic hazard and risk. There almost appears to be a radial response when moving out into the field from Eemskanaal onwards. The clusters within the South-West and OPP controls have a negative effect as well, but to a smaller degree. The next ring of controls tends to be relatively indifferent (Leermens, Siddeburen, South-Central), and moving further along the outer ring of controls is generally green. A shift in directionality can be observed between the south (control: South-Central) and the North (control: ZND) of the field depending on the objective function: the controls favour production from the more densely populated South of the field for hazard, and the sparsely populated North of the field for risk.

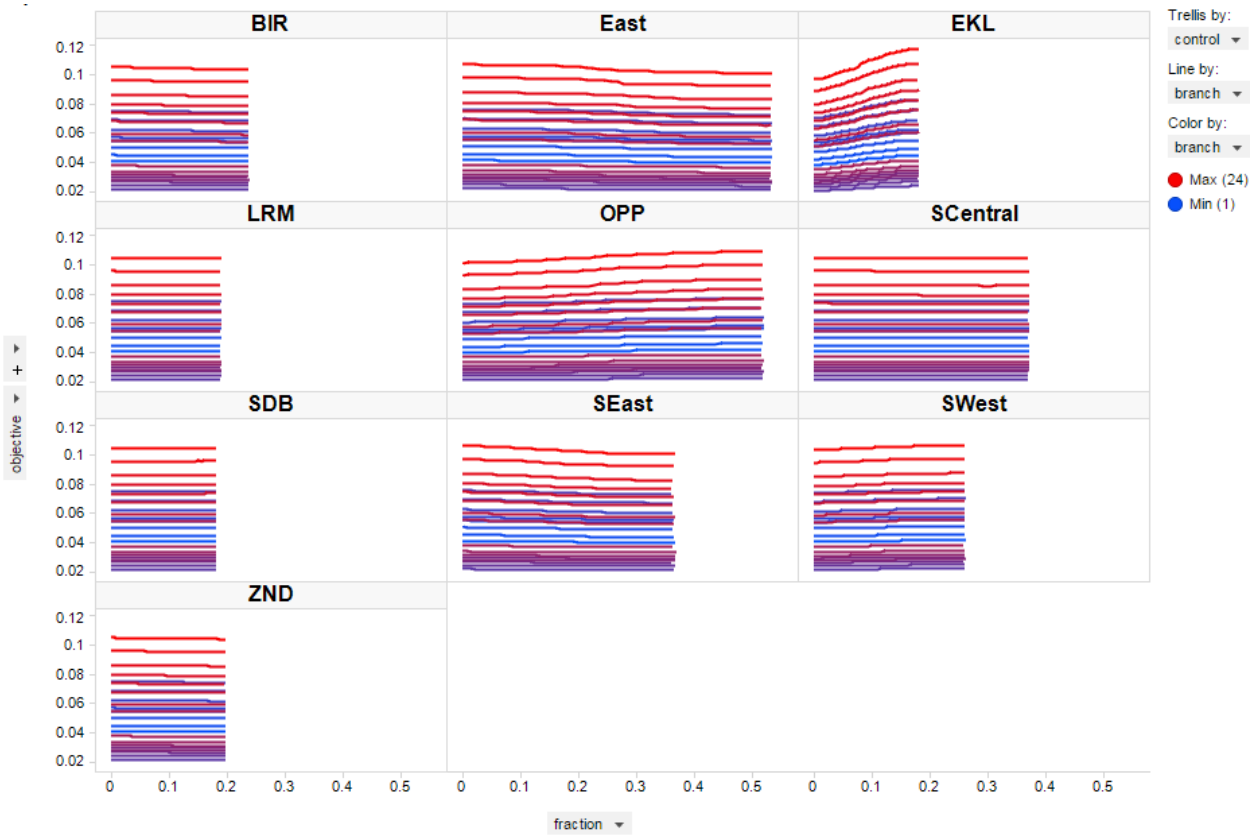


Figure 6-7: Partial dependence plots of pwPGV with respect to the production fraction of each control, for all 24 branches of the Hazard tree

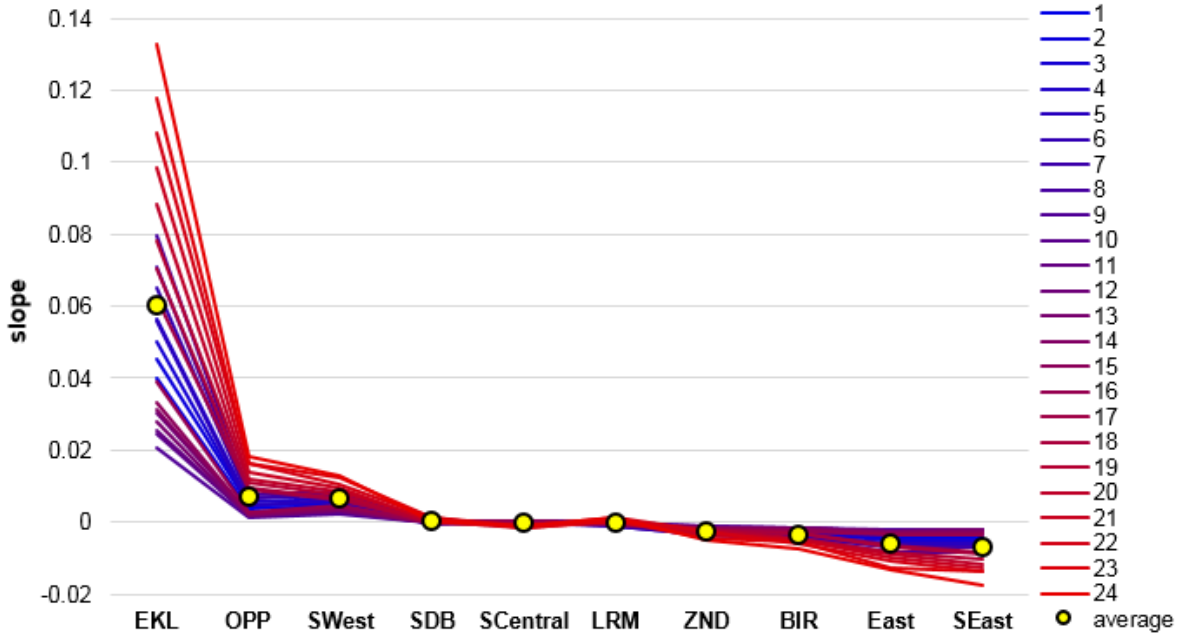
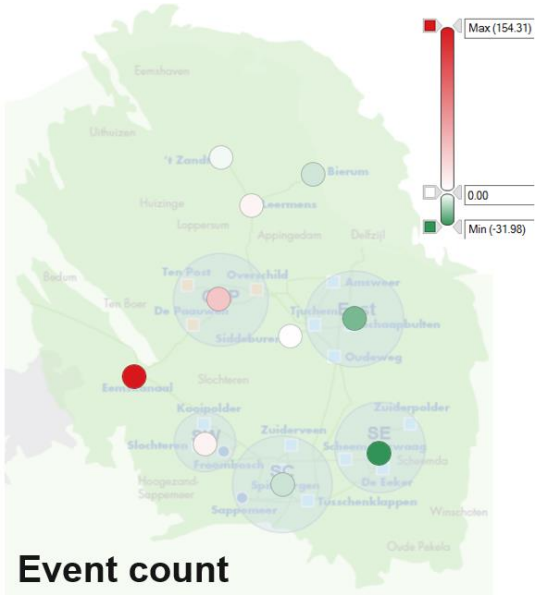
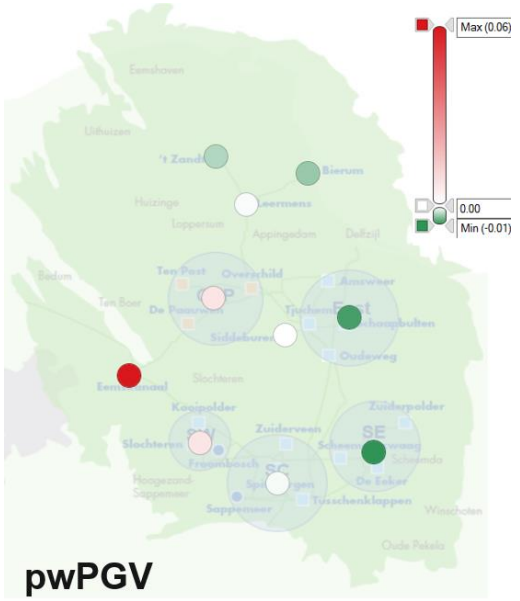
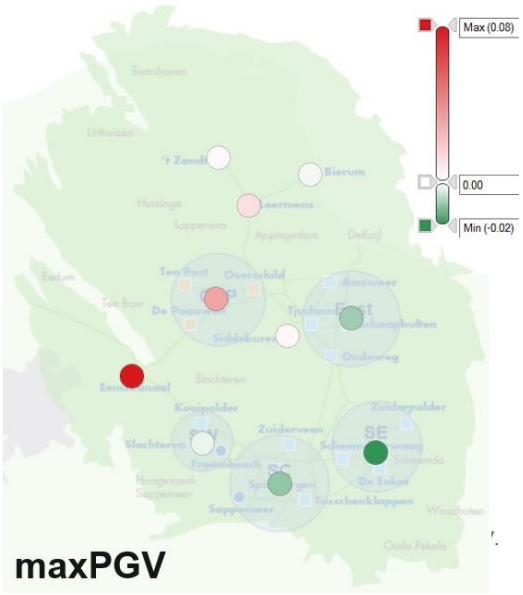
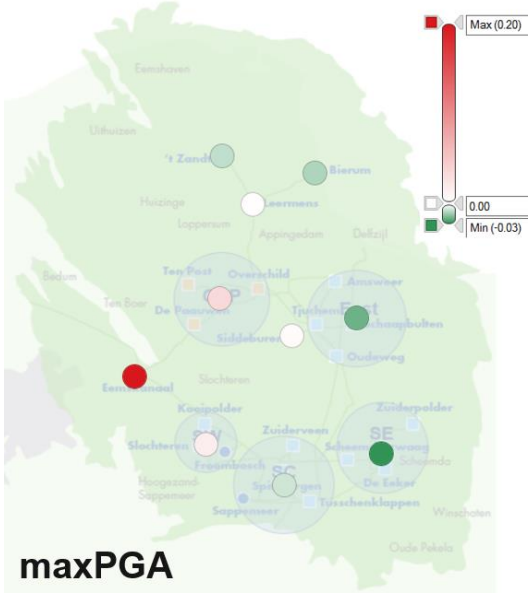


Figure 6-8: Best fit values of linear slope of the partial dependence plots for each control. All 24 branches of the Hazard tree are displayed, varying from blue (1) to 24 (red).









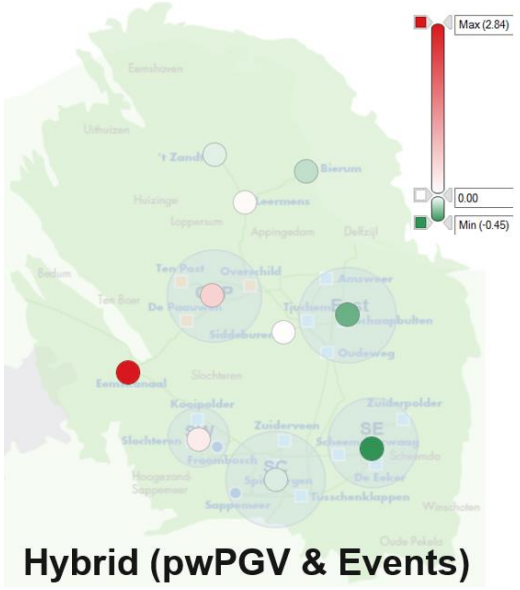


Figure 6-9: Averaged slopes from partial dependence analysis for all objective functions

7 Operationalization of results

7.1 More realistic reflection of operational envelope

The production optimisation study was consciously implemented such that the full control space could be tested. As described in section 3.4, each control was allowed to produce anywhere between zero and maximum capacity, whereby maximum capacity was imposed by means of as a simple THP constraint. In reality, the operational space is smaller. For instance, production clusters are bound by minimum throughput levels within the production facilities, and cannot physically produce at e.g. 1% of maximum capacity. The maximum capacity, in turn, is in reality not governed by a single tubing head pressure. The production rates that can maximally be achieved are governed, amongst others, by the compressor envelope, possible compressor speed restriction due to noise constraints, and the backpressure that is exerted on the compressor outlet by an interplay between the location of the production cluster in the ring, the number of transfer stations that are in use and their respective rates, the operational settings within the ring (e.g. whether or not the double lines can both be used, ring split), and the backpressure within the GTS network.

7.2 Model description

The importance of accurately modelling the surface facility constraints has long been realized by the NAM. Over the last 30 years this has evolved into the Gas Integrated Production System Model (IPSM) software toolsuite called GenREM, a robust, strongly integrated business tool supporting operational gas field development and forecasting. GenREM is in operation for all NAM operated gas fields in the Netherlands and has gained countless capabilities over the years, continuously evolving to address the changing technical and regulatory demands of both the gas fields and the stakeholders of NAM.

Development of GenREM has been a combined NAM/ORTEC effort between mathematicians, IT specialists and Petroleum engineers, resulting in a time-step driven gas capacity and production forecasting tool, using a detailed gas facility treatment network definition. This network solver can be coupled to the Mores subsurface reservoir simulator, resulting in an integrated model capturing the entire Groningen system up to the export points to GTS.

Modelling the Groningen Ring, as described in Appendix 3, is a challenge for a surface facilities model. For this purpose, a dedicated surface network ring solver has been developed. Being a pressure balanced solver, GenREM is completely thermal driven. Every cluster is modelled in detail, defining every physical Groningen surface device in place. All the cooling devices are based on Heat Transfer Research Institute (HTRI) modelling. In Figure 7-1 a typical layout of the devices associated with a cluster location is shown.

The Groningen specific design of the centrifugal compressors is completely implemented, up to the level of its efficiency, which is based on a neural network implementation based on actual conditions. The compressor characteristics are based on the envelopes and checked against anti-surge control and speed line limits. Recycling is fully implemented. Multi stage power control usage has been specifically designed for GenREM in close cooperation with the NAM and Siemens engineers (the supplier of the compressors).

GenREM handles constraints for numerous physical parameters like flow, pressure and gas quality throughout the whole surface definition to ensure realistic results of the simulation. Calibration of the input parameters is established by means of big data analysis of real-time production data.

GenREM is being used for short, medium and long term forecasting of gas production and capacity. Additionally, it is also being used for electricity forecasting due to the high power demand of the compressors.

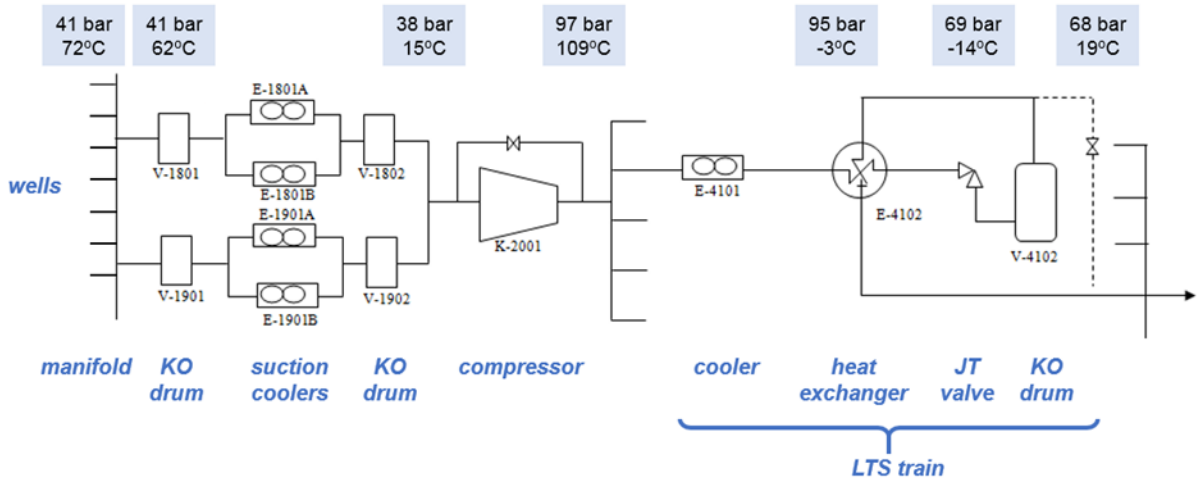


Figure 7-1: Typical layout of surface facilities at a cluster location

A high-level description of the operational restrictions which are captured in the surface network model is given in Appendix 3.

It should be realized that in the GenREM evaluation is based on the assumption that the full installed equipment is available at all times, and that no functionality deterioration takes place. Additionally, the restrictions and distribution of the gas across the custody transfer stations (OV's) is not addressed in this optimisation.

7.3 Production regions implemented in model

The 10 controls from the optimisation, Table 5-3, were further reduced to 5 production regions with the conditions that they are:

1. **Specific enough to implement the outcome from optimisation**: Although a single volume target is assigned to a production region, cluster prioritization within the region allows close adherence to the optimisation outcome.
2. **Large enough to allow for operational flexibility**: Allows for redundancy within each region group in case of operational upsets or planned maintenance.

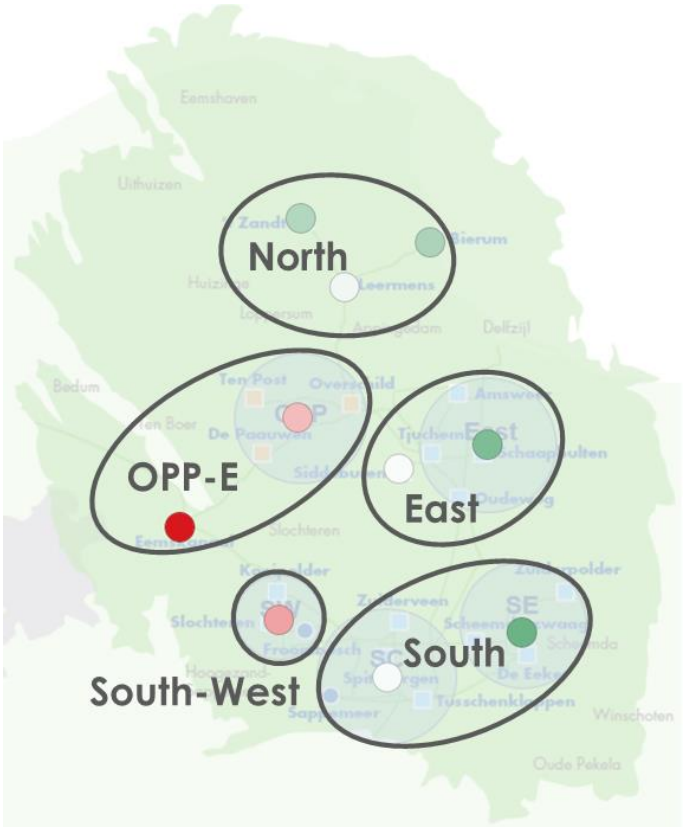


Figure 7-2: Production regions used for Genrem-Mores operational modelling

These are the same regions as defined in Figure 1-7 with the excetion that South-East and South-Central have been lumped into a single South region where the central clusters are operating at a lower priority. Each of the 3 regions to the East of the field (North, East and South) comprises both "green" and "white" clusters from the outcome of the population weighted partial dependence analysis (chapter 6). This is encoded in the GenREM -Mores modelling through the use of cluster priority. The volume target for a region is preferentially achieved through the production of "green" clusters. If insufficient capacity is available to fill a target volume, lower priority "white" clusters are used to achieve the target.

Table 7-1: Production regions and associated clusters used for Genrem-Mores operational modelling. Within a region, clusters marked green are preferentially produced before those marked white. This color coding corresponds to partial dependence analysis. Region capacity values are based on 2017 average modelled values.

Region	Clusters					Capacity	
						[BCM/YEAR]	[mln NM3/DAY]
North	BIR	ZND	LRM			11.8	32.4
East	AMR	SCB	TJM	OWG	SDB	16.8	46.0
OPP-E	OVS	POS	PAU	EKL		16.6	45.6
South (East - high priority)	EKR	SZW	ZPD			10.2	27.8
South (Central - low priority)	SPI	ZVN	TUS	SAP		11.3	30.9
South-West	SLO	FRB	KPD			7.7	21.1
						74.4 BCM/YR	204 mln NM3/DAY

Decreasing Priority

Additional modelling assumptions used are:

- Maintenance schedule as per 2017 Operating Plan
- Flat seasonal production (i.e. no variation in production between summer and winter)
- Conversion to 2nd stage compression starting in 2022 with 2 clusters per year in the first 3 years, followed by 1 per year.

The deterministic runs with the surface network model were done for a 10 year period to allow for review of the behavior beyond the 5 year optimisation window.

7.4 Implementation of hybrid optimisation outcome

Given that the hybrid optimisation can achieve close to the minimum both in terms of risk and nuisance, Table 5-4, this scenario is selected for implementation in the coupled Genrem-Mores model to assess its operational feasibility.

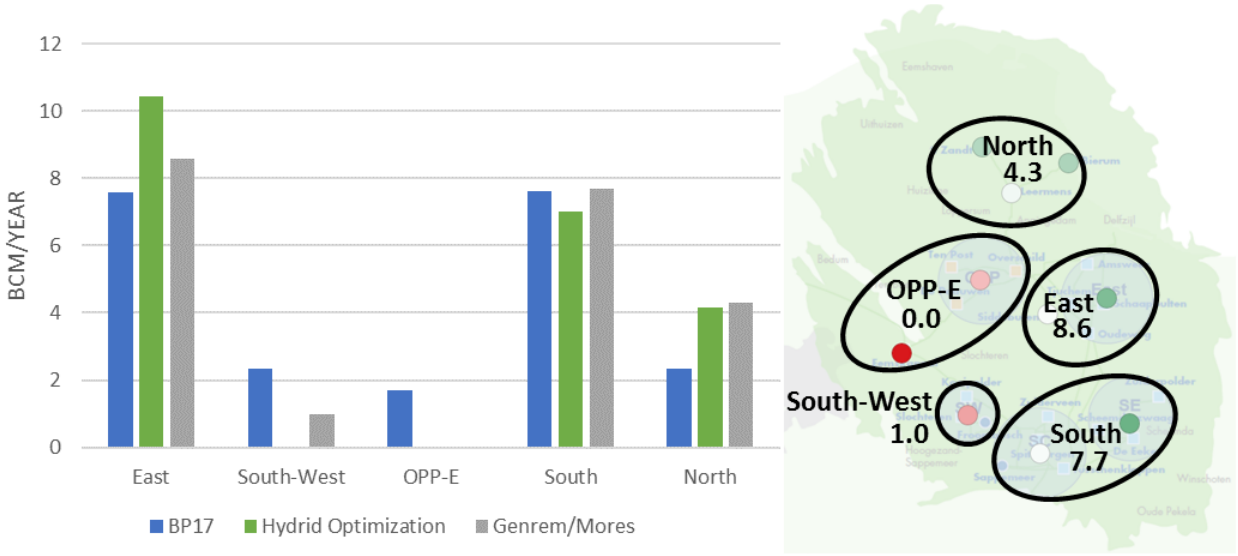


Figure 7-3: Default rate targets (Bcm/year) assigned to production regions in Genrem/Mores.

7.4.1 Operational considerations with regards to the South-West production region

Production from the south-west production region, that includes the clusters Slochteren, Froombosch and Kooipolder, require some special consideration with respect to operationalization. The NorGron pipeline ties-in the south-western part of the ring. Using gas from further afield clusters for injection into UGS Norg would require a more extensive ring segregation, and consequently closure of more custody transfer stations during injection period resulting in limited distribution options feeding the G-gas market.

Although the different optimisations, as well as the partial dependence analysis, suggests to reduce the production from the south-west clusters, it is proposed to apply a low start-up priority production philosophy for these clusters. The cluster capacity remains available, with an operational stand-by volume of 1 Bcm and a maximum of 2.6 Bcm (in an average year).

7.5 Results based hybrid optimisation

The results from the GenREM /Mores model are given in Table 7-2. This is based on the rate distribution from Figure 7-3. The reference values are also based on GenREM /Mores model, but with the areal rate distribution as per Business Plan 2017 (BP17) and an total field cap of 21.6 Bcm/year. The reference values are slightly different from those quoted in Table 5-4 due to the inclusion of detailed surface network model and the stochastic uncertainty between the models. Reduction in hazard metric values are close to those achieved by the unconstrained optimisation, 13 vs 16% reduction in population weighted PGV metric.

Table 7-2: Hazard and Risk based metrics from Genrem/Mores modelling. Reference values are based on a Genrem/Mores model using production distribution as per Business Plan 2017.

	HRA Metric Values			
	pwPGV	**Max PGA**	**Max PGV**	**Event count**
Reference (BP17)	**0.056**	**0.255**	**0.144**	**118**
Genrem/Mores	**0.049**	0.238	0.139	**101**
	-13%	-7%	-4%	**-14%**

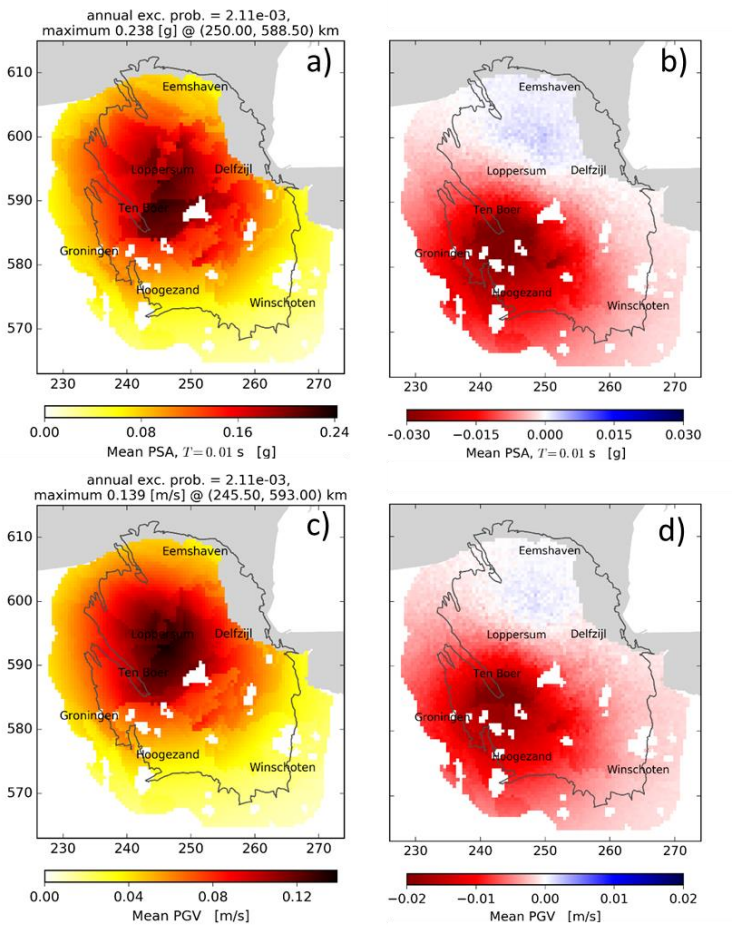


Figure 7-4: Mean PGA and PGV of the optimised case (a,c) and corresponding improvement compared to the BP17 Reference case (b,d).

The areal improvement in hazard for the optimised offtake distribution is shown in Figure 7-4 b. Compared to the hybrid optimisation case in Figure 5-11 the pattern of improvement is very similar, with the more densely populated south-west achieving the biggest gain.

The forecasted production for a 10 year period (2018 to 2027) is shown in Figure 7-5 a. The model predicts that the 21.6 N.Bcm annual production volume can be sustained for the entire period. The lower priority south-central clusters only contribute to the offtake volume as required to achieve volume target, a gradual ramp-up can be seen over time in Figure 7-5 b with ongoing depletion of the field (and subsequent reduction in capacity). The low priority cluster contribution to the North and East production regions are highlighted in Figure 7-6.

In terms of production capacity, Figure 7-7, it is clear that the 3 eastern regions (North, East and South) do initially have spare capacity that could be used to allocate production between the regions.

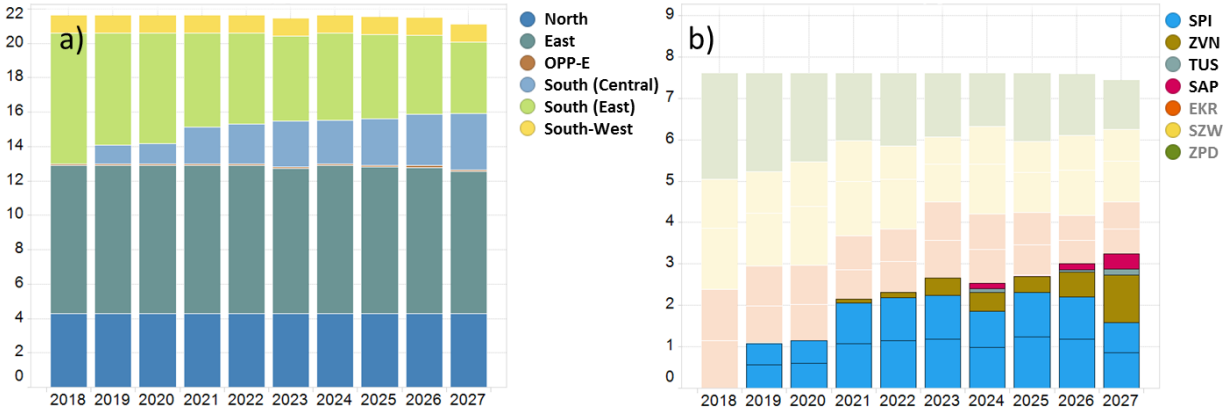


Figure 7-5: Production distribution for 10 year forecast using an optimised areal distribution a). For production region South the higher priority eastern clusters (EKR,SZW and ZPD) are preferentially produced to those of the central area (SPI,ZVN,TUS and SAP) b).

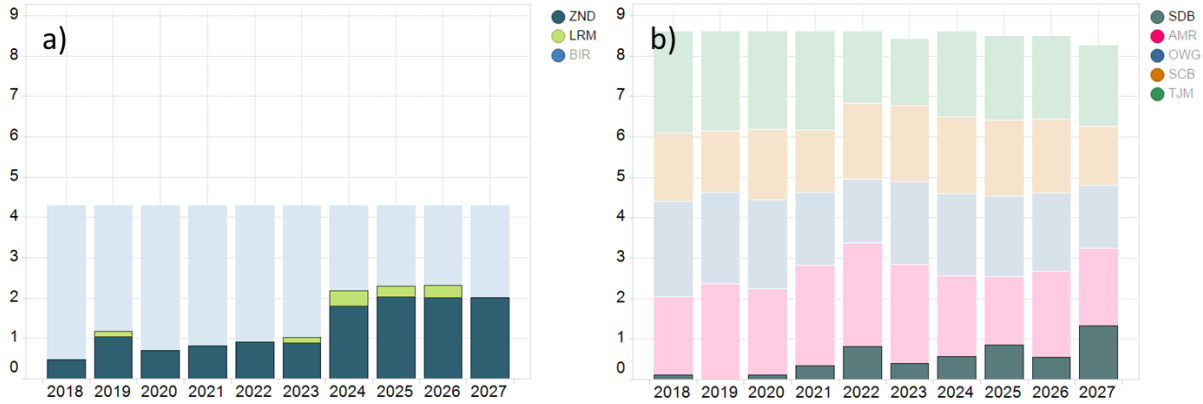


Figure 7-6: Contribution of lower priority (white) clusters to North (a) and East (b) production region.

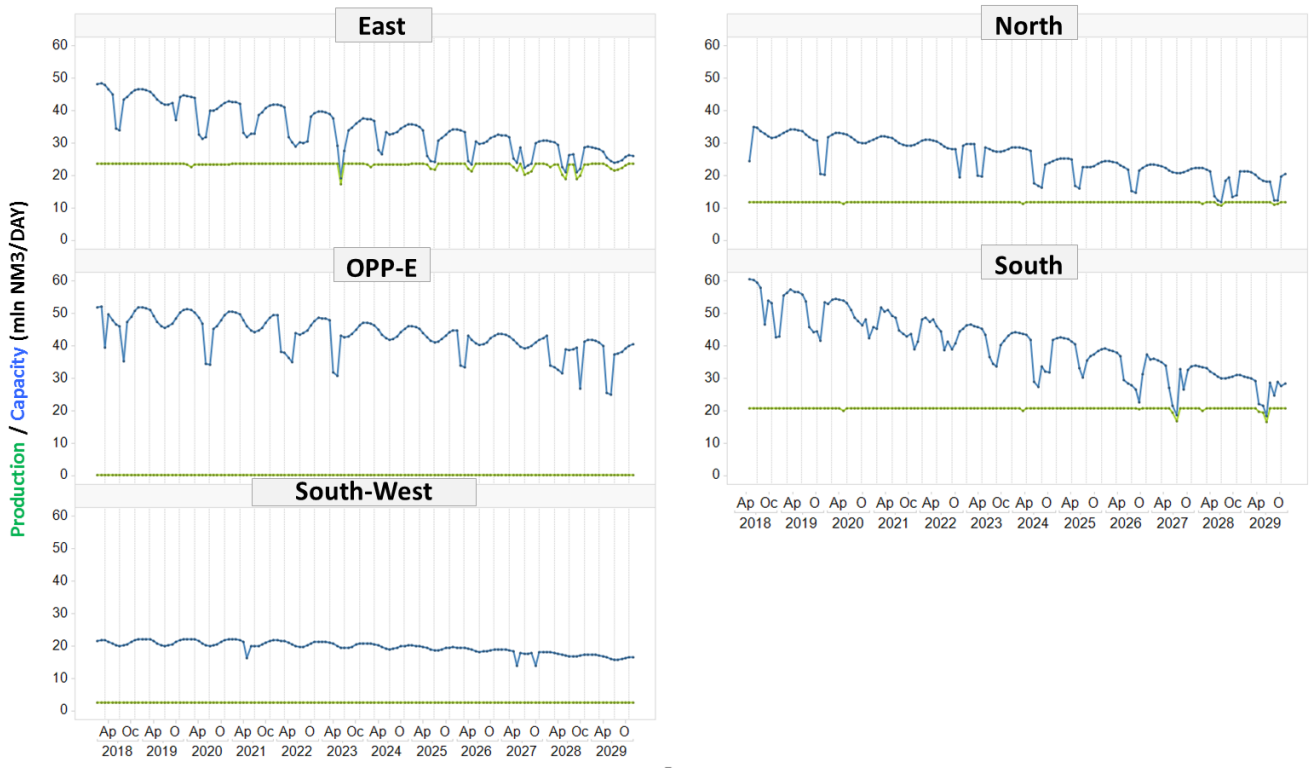


Figure 7-7: Production and capacity of the various production regions for the optimised 21.6 Bcm case

7.6 Impact of volume re-distribution between low seismicity regions

Although the optimisation does result in a single suggested distribution of production volumes, it is also clear from the partial dependence analysis that the three "green" regions (North, East and South) do act in a similar manner. Assessing the forecasted impact on seismicity from re-distributing volumes between these regions is interesting, both from an operational standpoint as well as the flexibility it provides with respect to the Meet- en Regelprotocol, see chapter 8 and Reference [19]. Three scenarios are investigated with a 20% volume swing between the regions, as indicated in Figure 7-8. In Table 7-3 the impact on the seismicity metrics are shown. The field wide average metrics are essentially the same given stochastics uncertainty. From the areal hazard plots there are, however, some apparent trends. With more volume allocated to the North region (Case 2) there is a resulting increase in hazard in that area. In Case 3, where more volume is allocated to the south, there is a corresponding reduction in volume offtake and hazard in the North, but also with some increase in population weighted PGV (over the entire field).

Permanent volume re-distribution between the regions does, however, have a significant impact on operational flexibility and the ability to maintain long-term production plateau. Although the overall field annual offtake of 21.6 Bcm can almost be maintained in all three cases, Figure 7-10, increasing volume offtake in the East or South region does cause those regions to be capacity constrained considerably earlier (Figure 7-11) compared to the base case as shown in Figure 7-7.

	North	East	South	SUM
Reference (BP17)	2.4	7.6	7.6	17.6
Base Case	4.3	8.6	7.7	20.6
1. East plus 20%	3.4	10.3	6.9	20.6
2. East minus 20%	5.2	6.9	8.5	20.6
3. South plus 20%	3.5	7.9	9.2	20.6

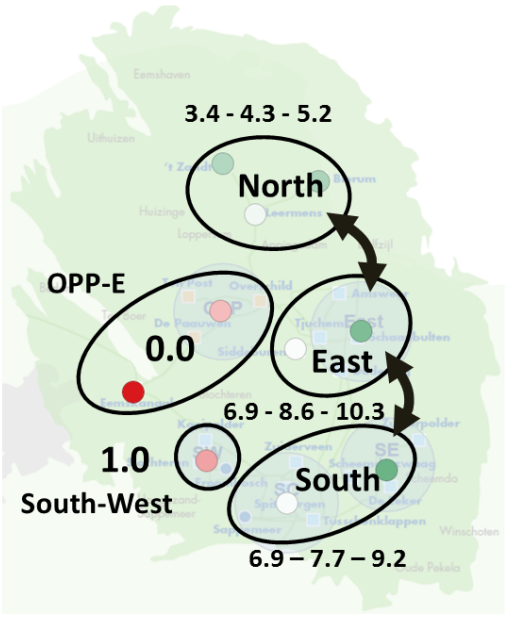


Figure 7-8: Volume distribution between North, East and South Production region.

Table 7-3: Impact on seismic metrics by redistributing production volumes between North, East and South production region.

	HRA Metric Values				HRA Improvement			
	pwPGV	Max PGA	Max PGV	Event count	pwPGV	Max PGA	Max PGV	Event count
Reference	0.056	0.255	0.144	118				
Base Case	0.049	0.238	0.139	101	-13%	-7%	-4%	-14%
1. East plus 20%	0.049	0.239	0.138	101	-13%	-6%	-4%	-14%
2. East minus 20%	0.049	0.235	0.137	102	-13%	-8%	-5%	-14%
3. South plus 20%	0.049	0.236	0.135	101	-12%	-7%	-6%	-15%

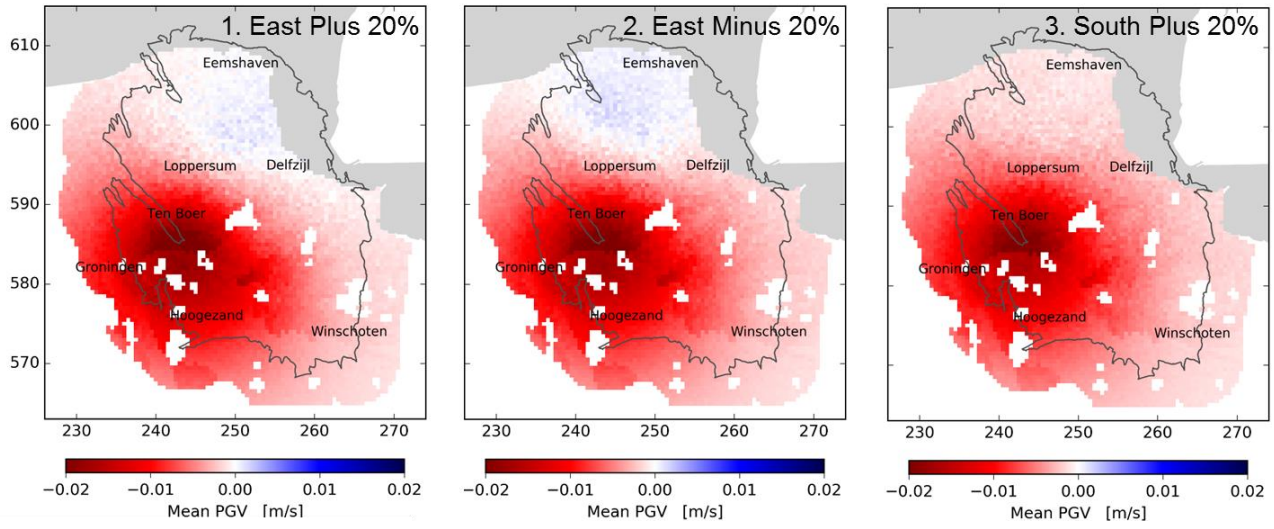


Figure 7-9: Areal impact of redistributing production volumes between North, East and South production region. Plots are differences with respect to the BP17 reference case.

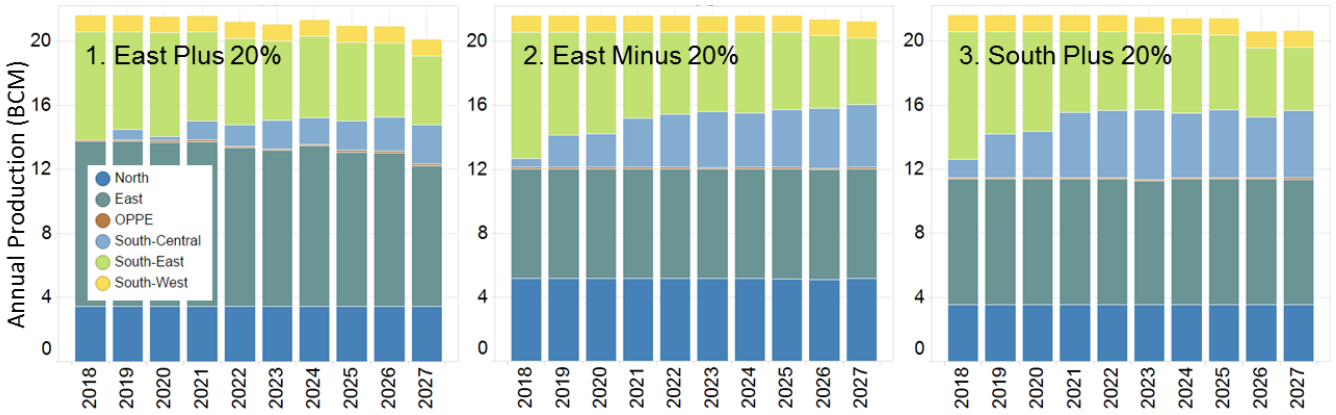


Figure 7-10: Impact on annual production from redistributing production volumes between North, East and South production regions.

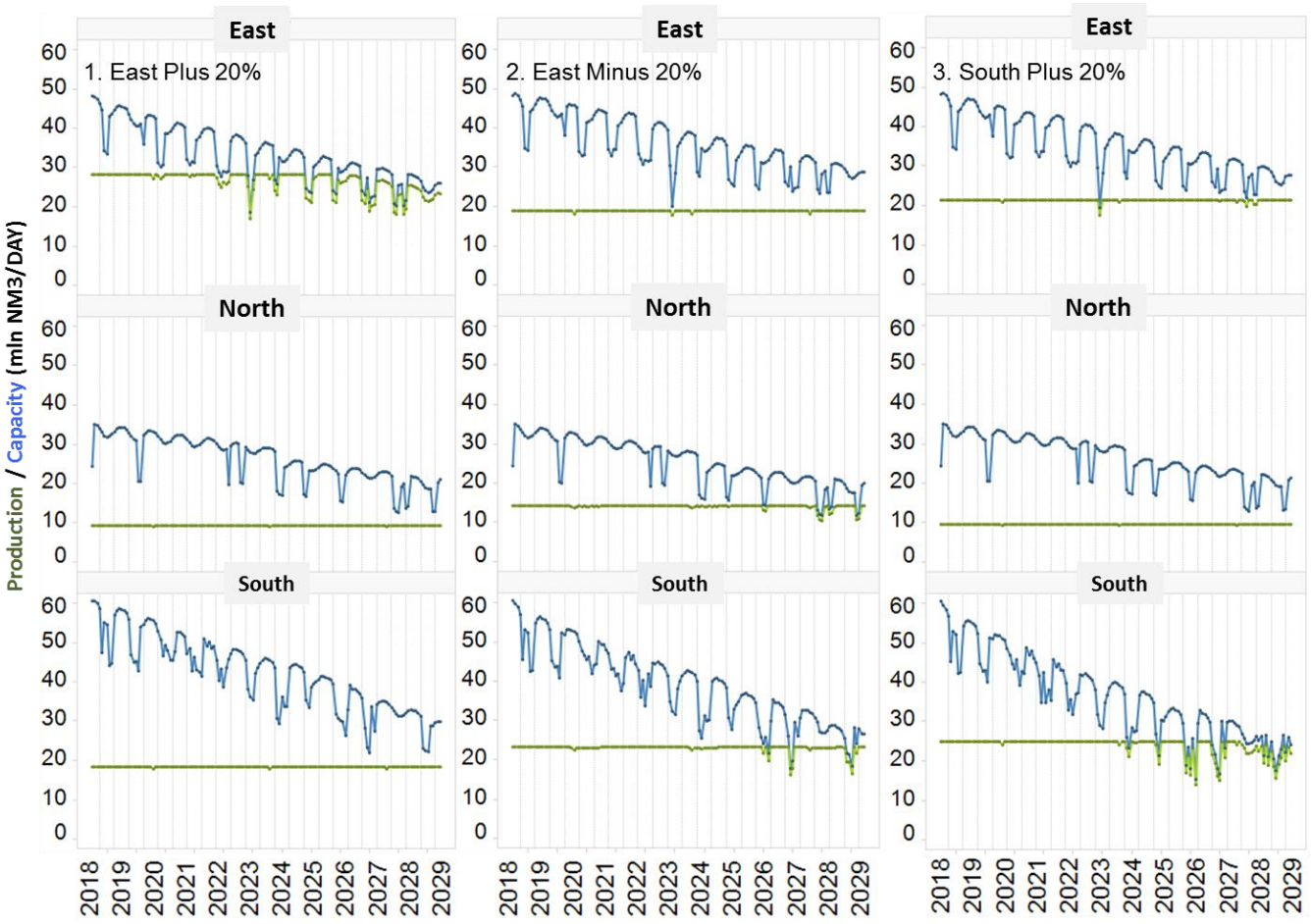


Figure 7-11: Impact on capacity from redistributing production volumes between North, East and South production regions.

7.7 Impact of volume re-distribution within regions

As mentioned in section 7.3, volume offtake within the three main regions is distributed following a priority list, where the "green" clusters are preferentially produced before the "white" clusters. However, operationally clusters must have a certain minimum production to ensure that their capacity is readily available in case of disruption to other clusters within the region (stand-by rate). Additionally, maintaining flat production within the region is more difficult in case certain clusters are operated at very low load factor.

The impact on seismicity from operating all clusters at equal priority within the region (i.e. the clusters are operated at equal load factor) is investigated for the South and North region. In these regions, the low priority clusters contribute a significant fraction of the overall capacity, as evident from Table 7-1 for the South region.

The impact on seismicity is given in Table 7-4. In the south, moving production towards the west does result in a deterioration of the field-wide seismicity metrics, both in terms of population weighted PGV as well as tremor count (case 1). However, compared to the reference case there is still an improvement of about 10 percent. Operating the clusters in the north at equal priority yields less deterioration of the field wide seismicity metrics with respect to the optimum (case 2). The areal distribution of hazard is shown in Figure 7-12. The pattern of hazard differences with respect to the reference case largely stays the same. For case 1, the magnitude (color intensity) in the south-west has become reduced due to less improvement of hazard. For case 2, there is slightly higher hazard in the north compared to both the reference and base cases.

Table 7-4: Impact on seismicity metrics by producing clusters within the South and North regions at equal priority.

	HRA Metric Values				HRA Improvement			
	pwPGV	Max PGA	Max PGV	Event count	pwPGV	Max PGA	Max PGV	Event count
Reference	**0.056**	**0.255**	**0.144**	**118**				
Base Case	**0.049**	0.238	0.139	**101**	**-13%**	-7%	-4%	**-14%**
1. Equal South Priority	**0.051**	0.246	0.141	**107**	**-9%**	-4%	-2%	**-10%**
2. Equal North Priority	**0.049**	0.240	0.140	**103**	**-12%**	-6%	-3%	**-13%**

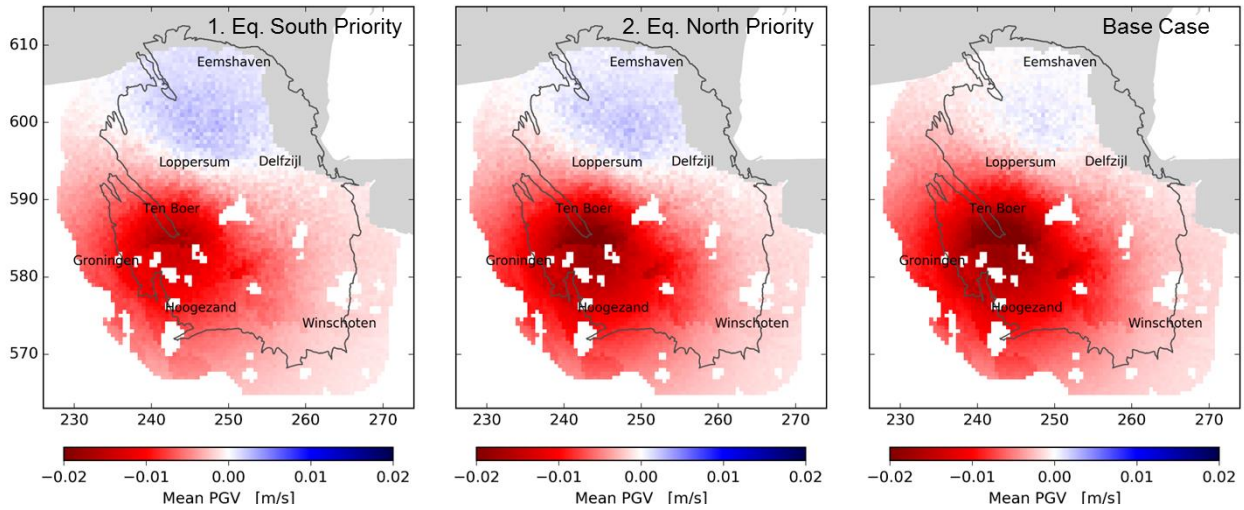


Figure 7-12: Impact on seismicity metrics by producing clusters within the South and North regions at equal priority. Plots are differences with respect to the BP17 reference case.

7.8 Impact of production from the South-West region

As mentioned in section 7.4.1, production from the South-West region is operationally important for the Groningen system. With the optimisation and partial dependence analysis suggesting to limit production from this area, the annual volume offtake from this area was notionally set to 1 Bcm. In this section, we investigate two additional scenarios. In Figure 7-13 the historical annual production from this region is shown. One scenario investigates the impact of maintaining offtake at this rate, 2.6 Bcm/y, which is close to the value as used in the reference case. The other scenario fully closes-in the production from the South-West region, as suggested by the hybrid optimisation. For both scenarios, these offtake changes in the South-West region are offset by opposite increase/decrease of production from the South region, thus maintaining the 21.6 Bcm total field cap.

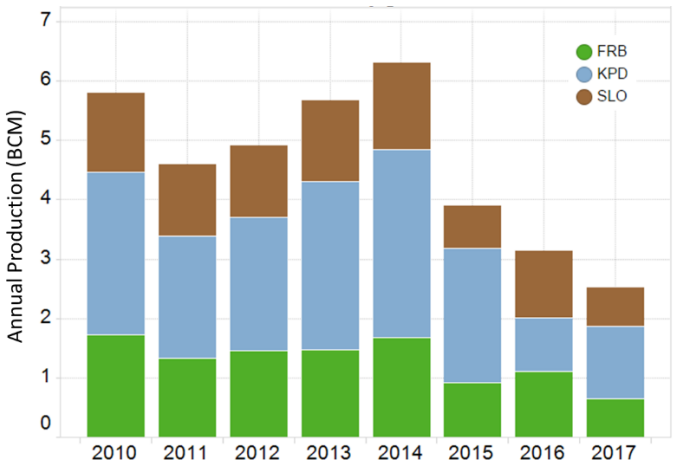


Figure 7-13: Historical annual production from the South-West region. 2017 values are an estimate as per the BP17 reference case.

The impact on seismicity is given in Table 7-5. Given the partial dependence analysis, it is not surprising that a higher annual production indeed does result in a deterioration of the seismicity metrics. It is worth noting that there still is a field-wide improvement compared to the reference case. The areal impact on hazard is shown in Figure 7-14. With an annual volume of 0 or 1 Bcm the results are, as expected, very similar. However, with a volume of 2.6 Bcm the range of improvement in PGV becomes considerably smaller.

Table 7-5: Impact on seismicity metrics by producing the South-West region at different annual volumes.

	HRA Metric Values				HRA Improvement			
	pwPGV	Max PGA	Max PGV	Event count	pwPGV	Max PGA	Max PGV	Event count
Reference	**0.056**	**0.255**	**0.144**	**118**	-	-	-	-
1. SW at 0.0 BCM/year	**0.049**	0.240	0.139	**100**	**-14%**	-6%	-4%	**-15%**
2. SW at 1.0 BCM/year	**0.049**	0.238	0.139	**101**	**-13%**	-7%	-4%	**-14%**
3. SW at 2.6 BCM/year	**0.052**	0.247	0.142	**110**	**-7%**	-3%	-1%	**-7%**

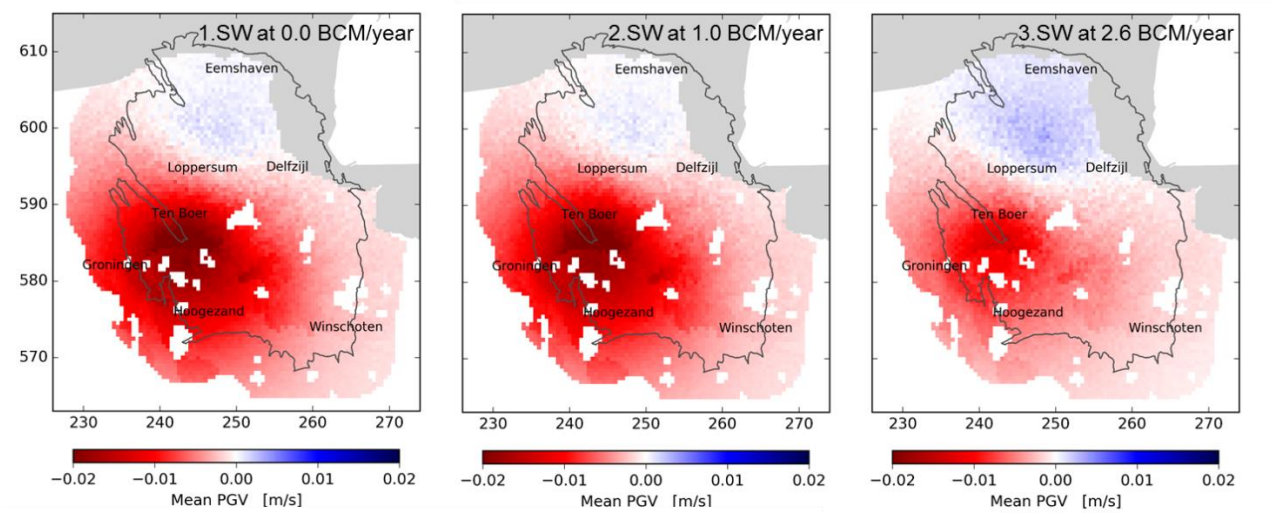


Figure 7-14: Impact on hazard by producing the South-West region at different annual volumes. Plots are differences with respect to the BP17 reference.

7.9 Comparing optimisation effort to reduction in overall field cap

In this section, the impact of regional offtake optimisation is compared to a reduction in the total field cap. This reduction in field cap was implemented by proportionally scaling the various regional volume caps, in line with the current regulatory framework. Both a 10 and a 20% reduction in cap were investigated, with respective annual production volumes of 19.4 and 17.3 Bcm. In Table 7-6 the impact on seismicity metrics are shown. Regional offtake optimisation performs better or equal to the case with 20% reduction in field cap (w.r.t. population weighted PGV and tremor count). In Figure 7-15 the areal difference in hazard is shown with respect to the regionally optimised 21.6 Bcm/year case. The pattern in hazard differences is similar in both cases. The optimised case improves hazard in the populous south-west whereas reduction in field cap improves hazard in the north-west compared to the optimised case.

Table 7-6: Impact on seismicity metrics by reduction in overall field cap (proportionally scaled reference case) compared to regionally optimised case.

	HRA Metric Values				HRA Improvement			
	pwPGV	Max PGA	Max PGV	Event count	pwPGV	Max PGA	Max PGV	Event count
Reference	**0.056**	**0.255**	**0.144**	**118**				
Optimized (21.6 BCM)	**0.049**	0.238	0.139	**101**	**-13%**	-7%	-4%	**-14%**
19.4 BCM CAP	**0.054**	0.246	0.140	**110**	**-3%**	-3%	-3%	**-7%**
17.3 BCM CAP	**0.052**	0.235	0.135	**101**	**-7%**	-8%	-6%	**-14%**

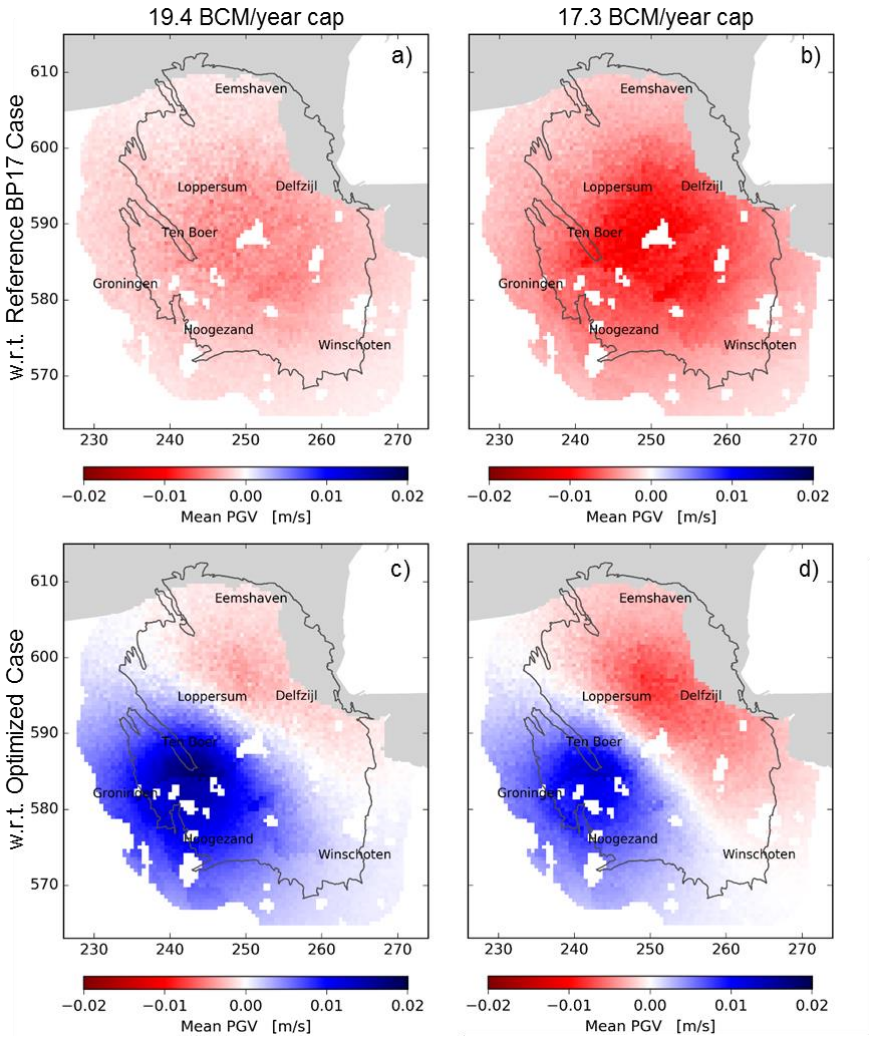


Figure 7-15: Difference maps showing the impact of reducing the overall field cap on hazard, for reductions of 10 and 20% with respect to the 21.6 Bcm/year. First row of plots (a-b) are with respect to BP17 reference case while second row (c-d) are with respect to the regionally optimised 21.6 Bcm case.

7.10 Implications

7.10.1 Operational feasibility of caps

In general, any increase in the number of (regional) caps will make it more difficult to operationally adhere to these caps. In case regional production caps are defined such that they (proportionally) sum up to the total year volume, the operating space will definitely narrow down towards the end of the production year, which effectively means a severe limitation throughout the entire year (point landing).

To increase the chance of success it is suggested to:

- Maintain a limited number of regions (as outlined in section 7.3);
- Introduce an operational band around the regional caps, and define the regional caps such that they add up to more than the total field cap.

Based on a qualitative analysis, it is expected that regional volumes can be achieved within a 20% bandwidth. Without such a bandwidth, it would be difficult to compensate for unexpected events, possibly resulting in an under-delivery of the volume required for security of supply at a flat production profile.

The control systems in Distributed Control System (DCS) may need adaptations to cope with new, modified or additional caps.

7.10.2 Maintenance strategy

Execution of maintenance involves the shut-down of production clusters. Those will impact the regional volumes that can be achieved, in particular when clusters are running at or near their maximum capacity and there is little flexibility in production rates. Hence the Long-Term Shutdown Plan (which captures all scheduled maintenance) will have to take into account the regional caps. Potentially the plan needs to be reviewed, for instance by considering for the southern clusters to start scheduling shut-downs in winter, to allow for more capacity being available in summer for injection into Norg UGS. This would be a complete reversal of the maintenance philosophy as executed over the past decades. Given that the Long-Term Shutdown Plan is a long-standing, rolling plan which integrates the deployment of thousands of man-hours from a multitude of departments and contractors, such a change would be a highly involved process.

Unexpected trips or failures at production clusters can have an impact on the ability to produce the regional volumes. For regions with a limited number of clusters the relative impact will be especially large. Potentially, the maintenance strategy for such regions may have to be adapted, with more focus on preventative maintenance.

Note that any change with respect to the current (BP17) production plan will impact the required timing for converting production clusters to second stage compression. Because of the (significant) shut-down times involved, the installation of second stage will significantly impact the ability to achieve regional volume targets. This will also have to be carefully considered in the Long-Term Shutdown Plan.

7.10.3 Norg UGS injection

The Norg underground gas storage acts both as a volume shifter to keep the Groningen gas production flat over summer and winter, and as a highly flexible capacity provider to meet high market demand on cold winter days, which is important for security of supply.

The maximum working volume that can be achieved is governed by many variables, including segregating the Groningen ring into high and low-pressure sections during parts of the injection season (Appendix E). With ring segregation the full working volume can be provided. However, for the optimised production scenario only a limited number of Custody Transfer Station can be kept open while the ring is segregated, resulting in limited flexibility in delivering gas to the market.

7.10.4 Electricity consumption

One implication of the production optimisation reported here, is that the production clusters in the South and East will run at high load factors. Given that all clusters are on compression, this involves maximizing the compression power that is consumed by those clusters. Hence the optimisation for seismicity involves a non-optimal deployment of the production facilities from an energy efficiency stand-point.

7.11 Opening-up clusters

The optimised production distribution for certain objective functions involves the reinstatement of the Leermens and/or ‘t Zandt production clusters. Reservoir modelling suggests that they have limited connectivity to the higher seismicity Loppersum area, as evident from the streamline plot in Figure 1-10 b). The hydrocarbon thickness map in Figure 7-16 clearly brings out the graben structure to the west of the ‘t Zandt and Leermens clusters that is associated with large fault throws. As a consequence, there is very limited gas-gas juxtaposition, and hence the east-west gas migration path is severely restricted. This dampens the impact from gas production by these clusters in the Loppersum area (section 1.5).

These observations are confirmed by the dynamic reservoir simulation model (Reference [8]). An important validation of the dynamic model is the addition of some 100,000 reservoir pressure measurements through enhanced use of THP data, Reference [20]. This has further improved the confidence in the model.

Re-opening the Leermens and ‘t Zandt clusters would involve a material change to the current status quo, and such a change should be carefully monitored. It is recommended that the resulting reservoir surveillance requirements are captured in the Study and Data Acquisition Plan. It is envisaged that this could involve an interference test with one the nearby Zeerijp wells, to confirm the (lack of) East-West reservoir connectivity by means of a dedicated pressure test.

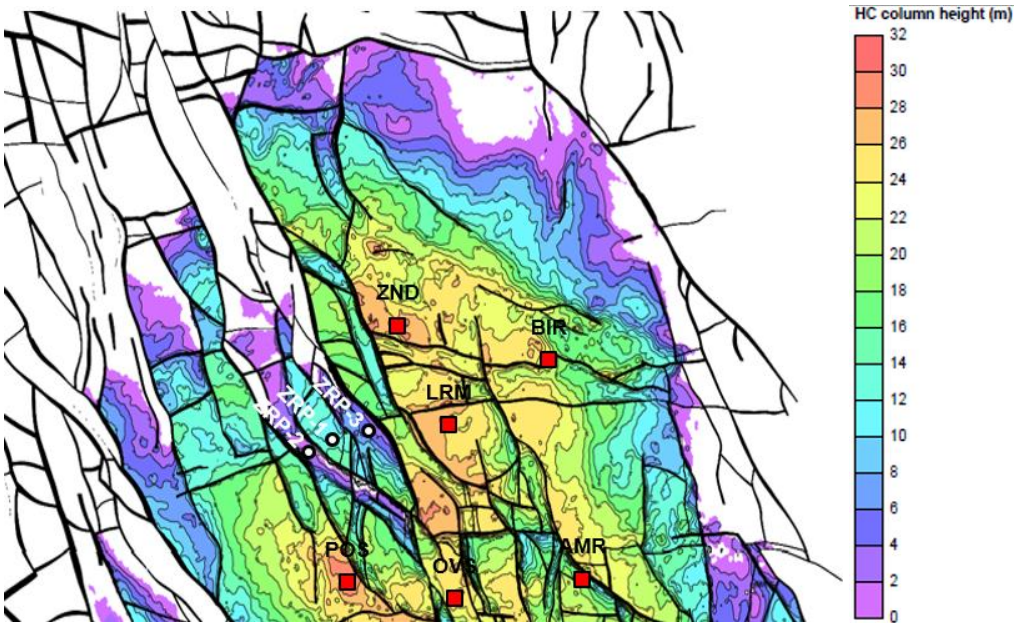


Figure 7-16: Hydrocarbon column thickness around the production clusters 't Zandt and Leermens

8 Interplay production optimisation and Meet- en Regelprotocol

The Production Optimisation study work may interact in two main ways with the Meet en Regelprotocol[11]: 1) In providing guidance to the choice of measures to be taken in case a threshold for a parameter has been exceeded and 2) in interpreting certain trends in parameters. The production optimisation work allows for additional insights to be developed into the understanding of local development in seismicity (section 0).

The Meet en Regelprotocol describes a three-tier system for control over the seismicity of the Groningen field. Relevant parameters are monitored in the field. For each of the five parameters monitored in the field as part of the Meet en Regelprotocol, there are three tiers of threshold values defined.

When a (combination of) threshold value(s) is exceeded due to a seismic event, this may trigger actions in the field, depending on the (combination and) tier of exceedance. In total 5 parameters are monitored:

- Three parameters can be exceeded by a single event without any preceding trend; highest measured PGA and PGV and occurrence of DS2
- One can be tracked over times as an indication for the seismicity over the full field; number of earthquakes with magnitude larger than M ≥ 1.5 during the previous 12 months.
- One can be tracked over times as an indication for the seismically most active area; earthquake density (M ≥ 1.0) over the previous 12 months.

"Earthquake-density" is important for an optimisation of production with respect to risk. The optimisation with respect to risk will preferentially move the hazard to areas of low exposure. It will for instance preferentially tolerate large hazard below the Eems estuary and other areas where relatively population density is low or buildings are relatively strong (able to withstand seismic loading). The optimisation could therefore cause a higher seismic rate and/or a higher chance of exceedance of a threshold value of the earthquake density over the previous 12 months in a low exposure area (whilst still reducing overall risk). In a similar vein, the production optimisation insights may provide also steer a decision to decrease production in area where a higher PGV (and/or PGA) will result from an earthquake.

From the above, it is clear that this study on Production Optimisation may provide important insights for operating the Meet en Regelprotocol. However, it should also be stressed that the Meet en Regelprotocol has been designed to work (somewhat) independently from the model frame also underlying the production optimisation work and that in some cases judgement will need to be used in order to suggest measures using insights from the Production Optimisation work.

The volume distribution as suggested in Chapter 7.4 maintains a balance across the 3 eastern production regions (South, East and North) that minimizes risk while allowing the annual 21.6 Bcm offtake to be maintained over the optimisation period. Additionally, from Chapter 7.6 it is evident that there is quite some flexibility in these volume targets. With a 20% shift in volume between the regions, the total field production can still be maintained with only a minor impact on the HRA metric values. In case of Meet en

[11] The Meet- en Regelprotocol is the Measurement and Control Protocol for the Groningen field.

Regelprotocol threshold exceedance (e.g. earthquake density), one possible intervention could thus be to temporarily shift volume away from the region experiencing exceedance.

9 Production fluctuations

9.1 Background

The updated Article 3.2 of Instemmingsbesluit WP2016 included an additional requirement for NAM to evaluate the ability to optimise cluster operations in order to limit (regional) production fluctuations.

> **Artikel 3**
>
> 2. De Nederlandse Aardolie Maatschappij B.V. onderzoekt of een alternatieve verdeling van de productie over alle regio's tot een lagere seismische dreiging of seismisch risico leidt en brengt daarover uiterlijk op 1 december 2017 ten genoegen van de inspecteur-generaal der mijnen, een rapport uit aan de Minister van Economische Zaken. Daarbij wordt tevens in detail voor alle clusters onderzocht of clusteroperaties verder geoptimaliseerd kunnen worden ten einde (ook regionale) fluctuaties zo veel mogelijk te beperken.

9.2 Temporal and spatial granularity

To address this request, it is important to realize that fluctuations of production occur at various temporal and spatial granularity, which have varying levels of controllability. Figure 9-1 provides a framework. Directionally, the top left diagonal represents an area over which there is no operational control.

- There are some 250 active production wells, most of which are some 50-60 years old. These wells are closely monitored for well integrity issues due to wear over time. On average, some 80 well interventions are needed annually to keep them in a good working order. These well interventions are typically resolved within days. On occasion a well encounters a more severe integrity problem which may require a work-over (months), or will trigger suspension or abandonment of the well (years or indefinite).
- There are on average some 400 cluster trips per year of which based on historical data 80% is resolved within hours. On occasion, "non-routine" root causes will require longer response times.
- Shutdowns of production clusters for regular/scheduled maintenance are in the order of 2-9 weeks
- Re-staging of a production compressor requires months

Hence high frequency production fluctuations (hour/day/week) at a low spatial level (well/cluster) cannot be avoided operationally. However, with increasing spatial and temporal granularity, these fluctuations can be averaged out to stable levels. In section 9.4 it is demonstrated that high frequency fluctuations (of small areal size) have a limited impact on reservoir seismicity: due to the physical nature of a compressible fluid in a porous medium, there is quick dampening of a pressure pulse.

Regardless whether production fluctuations may induce additional earthquakes, and whether or not there are creep effects in the reservoir, there is an increasing causality with seismicity when moving from the bottom-left quadrant towards the to-right quadrant. Prolonged production (year/decade) results in pressure depletion, which sets in motion the cause-and-effect chain as described in section 2.1, ultimately leading to seismicity. This is indicated as a blue arrow in Figure 9-1.

The seismicity models as used in this study relate cumulative reservoir compaction to seismicity, and do not reflect an impact of production fluctuations. In other words, the HRA model does not claim a benefit for flat production. For operating the field in reality, a precautionary principle applies; flat production should be implemented until proven otherwise.

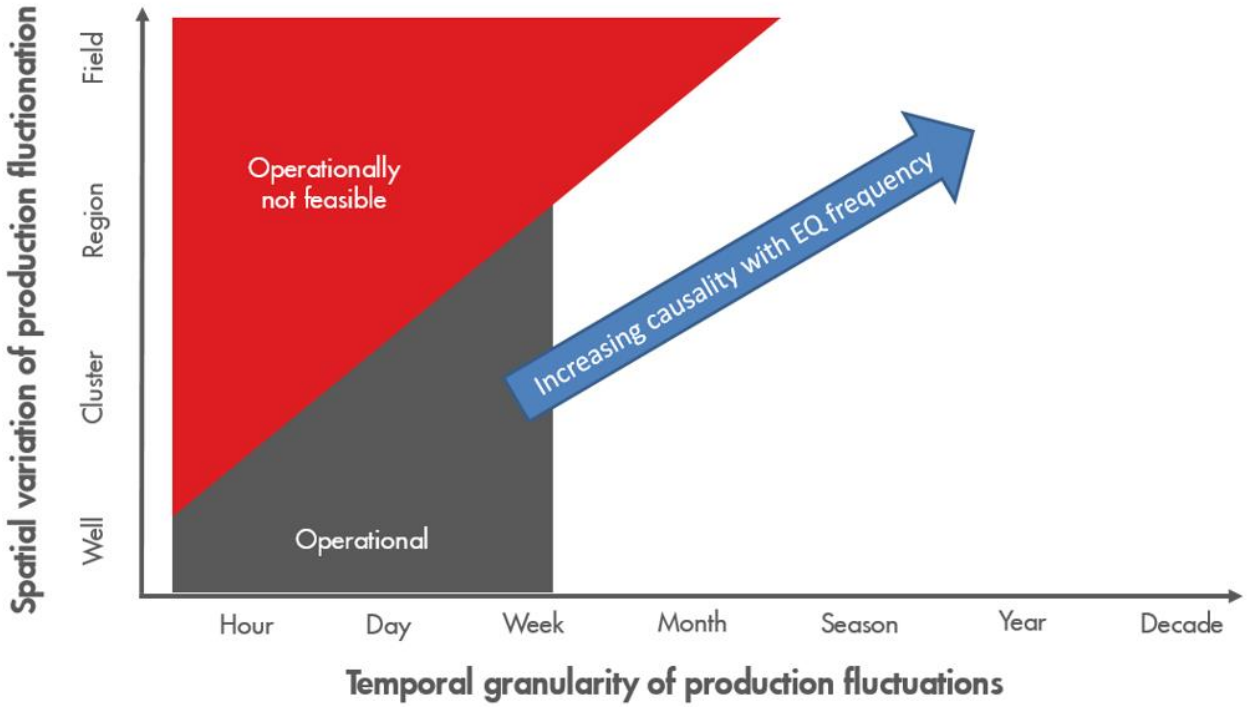


Figure 9-1: Spatial and temporal granularity of production variations in relation to earthquakes, indicating the operational domain.

9.3 Earthquake density in relation to pressure transients

When analysing pressure derivatives across the field, it was evident that the most severe pressure transients are associated with cluster shut-downs and start-ups. The analysis was based on actual historic data (on a well-by-well level), using daily timesteps. When running the dynamic reservoir model for daily time-steps, operational trips (especially those resulted in a full day of downtime) show significant pressure gradients. Potentially, the analysis could be repeated based on hourly production history to capture an even more complete picture of the true pressure transients. Such an analysis should however account for the fact that a production simulator does not reflect after-flow (or wellbore storage) effects, which will artificially amplify pressure transients in the immediate vicinity of the wellbore. However, these have a small areal extend.

When looking at the history match for the Loppersum area between 1/1/2016 and 31/7/2017 using daily timesteps, it can be observed that the fluctuations in pressure gradients are relatively modest, in line with

low offtake rates and relatively stable production of the regional production clusters. Nevertheless, the area was relatively seismogenic in this period, Figure 9-2. Similarly in the Eemskanaal area, pressure transients were virtually absent, and still earthquakes were triggered, Figure 9-3.

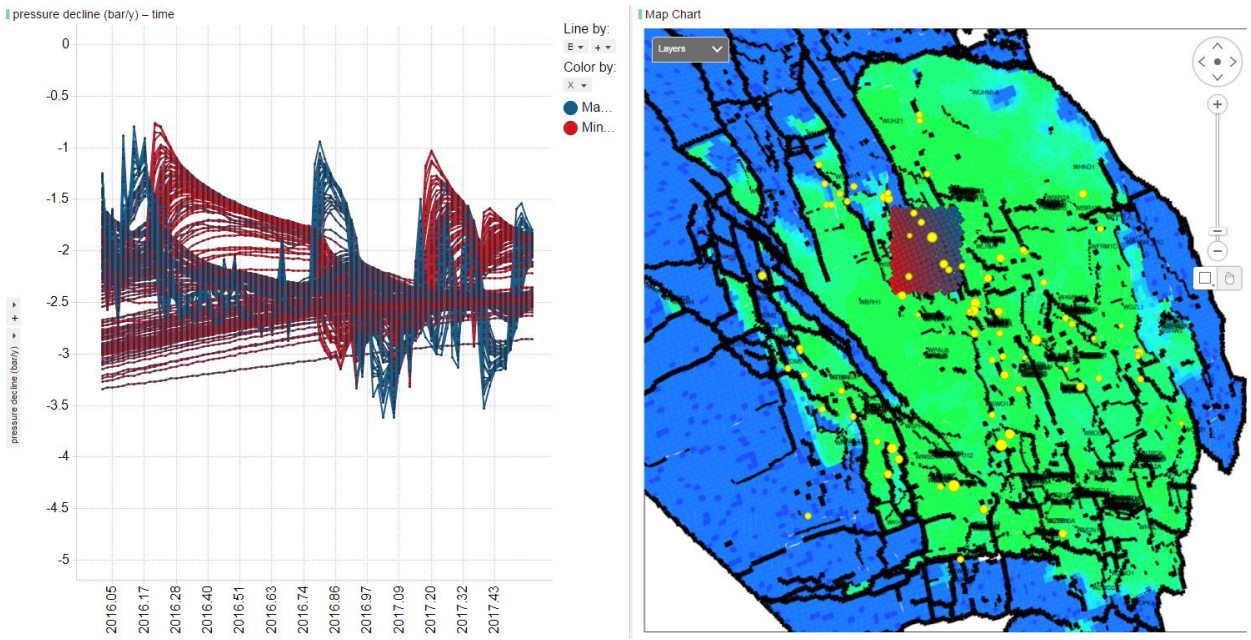


Figure 9-2: Pressure behavior for selected cells (red-blue squares) in the Loppersum area between 1/1/2016 and 31/7/2017, with reference to the earthquakes (M>1.0) in that same time window (yellow bubbles).

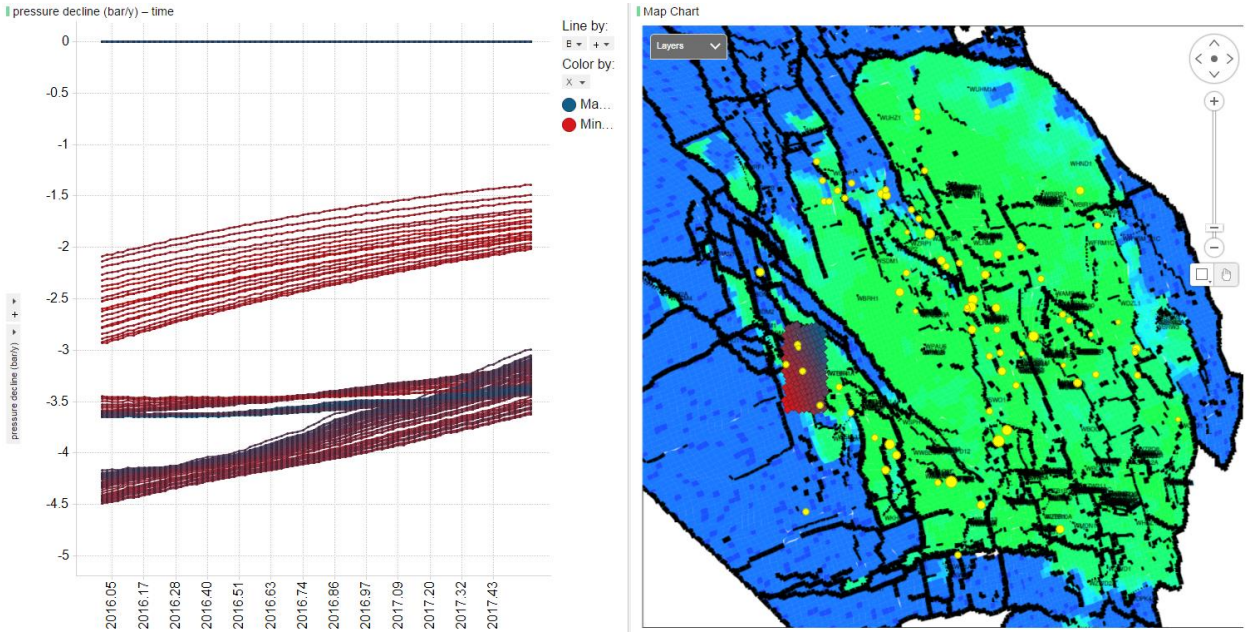


Figure 9-3: Pressure behavior for selected cells (red-blue squares) in the Eemskanaal area between 1/1/2016 and 31/7/2017, with reference to the earthquakes (M>1.0) in that same time window (yellow bubbles).

9.4 Lateral impact of production fluctuations (historic)

The simulation setup as described in section 9.3 clearly highlights the dampening of pressure transients away from the production wells. Figure 9-4 gives two representative snapshots of pressure transients; the LOPPZ and EKL area are virtually undisturbed, while the South and East show localized pressure transients around production clusters. Figure 9-5 provides an overview of the pressure transients in time for a selection of gridblocks, showing the diminishing transients while moving away from the Bierum cluster.

22/4/2016 19/8/2016

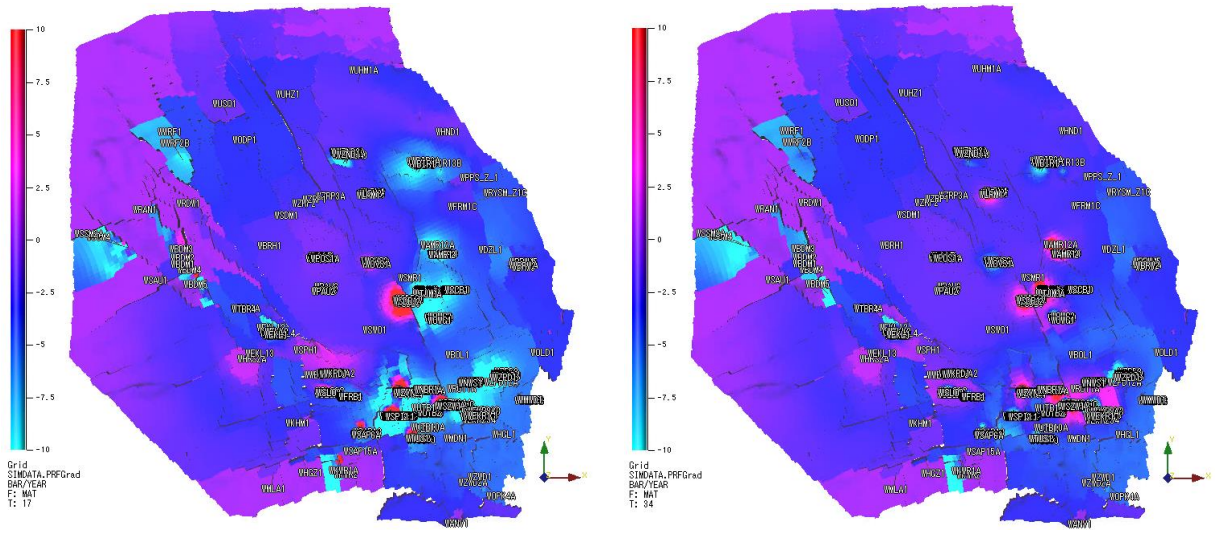


Figure 9-4: Snapshots of pressure transients (bar/y) in a map view.

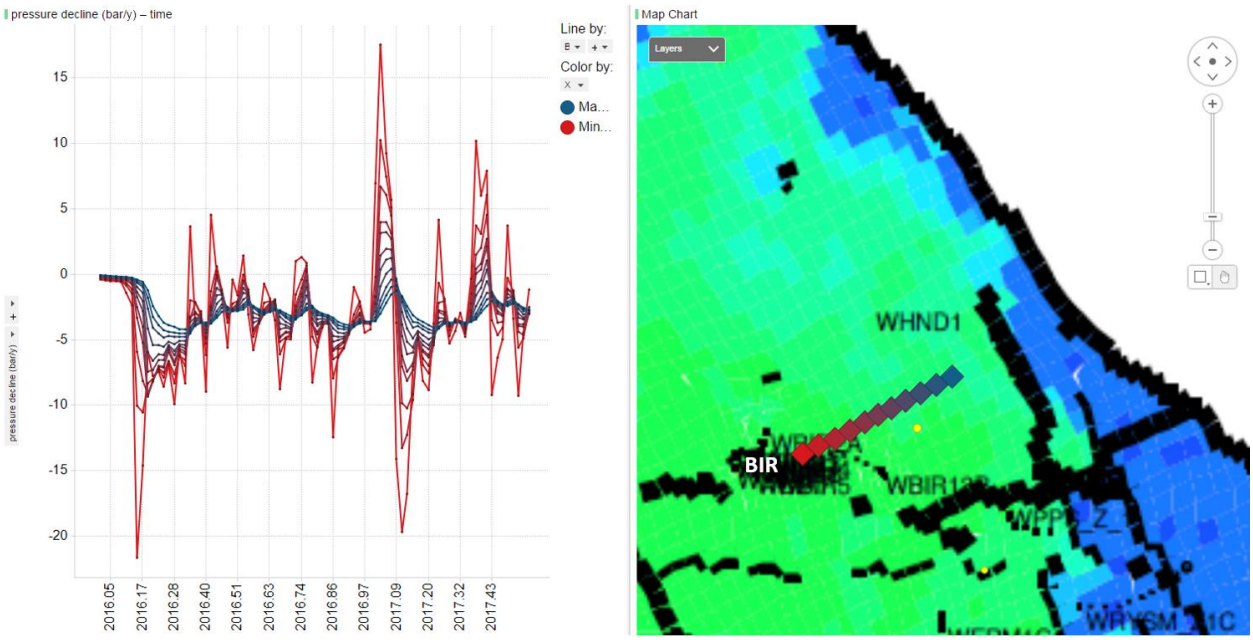


Figure 9-5: Dampening of the pressure transients away from the production cluster, example for Bierum. Color coded by gridblock from red (closer to cluster) to blue (further out).

9.5 Lateral impact of production fluctuations (generalized)

In order to generalize the transient behavior, the lateral impact of various temporal production pulses was tested with Pressure Transient Analysis software (Ecrin). A numerical model was setup for average Groningen properties. The impact of a series of five subsequent production pulses was tested at increasing distance away from the emitting point, varying from 50, 100, 200 to 400 m (Figure 9-6). The production pulses were constructed as an intermittent series of production (5 mln m^3/d) and shut-in of equal duration. The analysis was done for pulses of 6 hour, 1 day, and 1 month. Note that in this model setup an entire production cluster is represented by a single producing well. In reality, a production cluster involves between 6 and 12 wells which are spaced at least 50m apart. Hence this “point source” approach reflects a worst case.

Figure 9-7 clearly demonstrate the physical nature of a compressible fluid in a porous medium. A pressure pulse is quick dampened away from its emitting point. Intra-day pulses are monotonically declining at 200m away from the well. Daily pulses are so at 400m away.

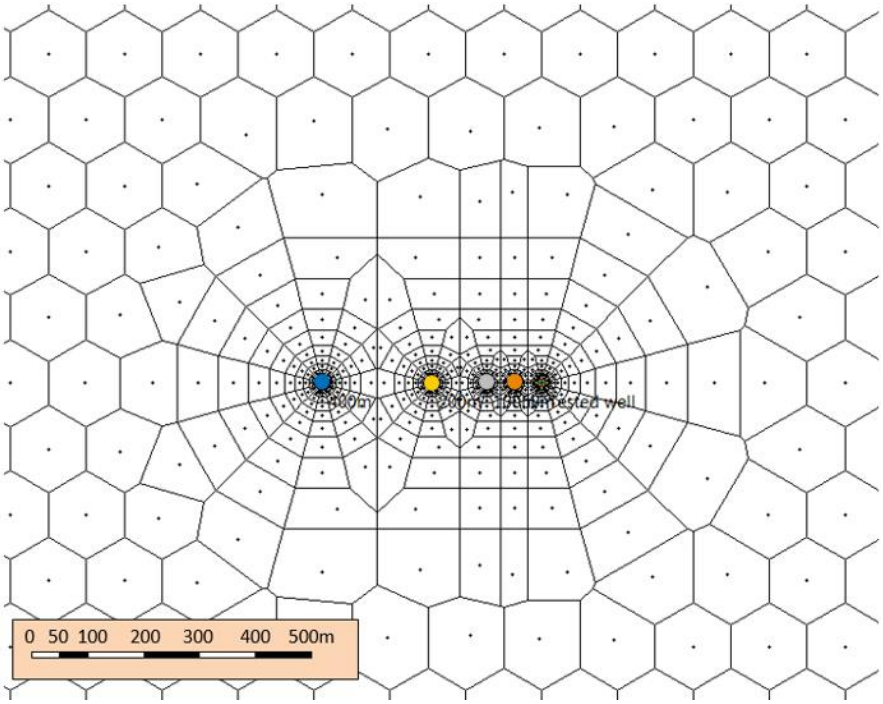
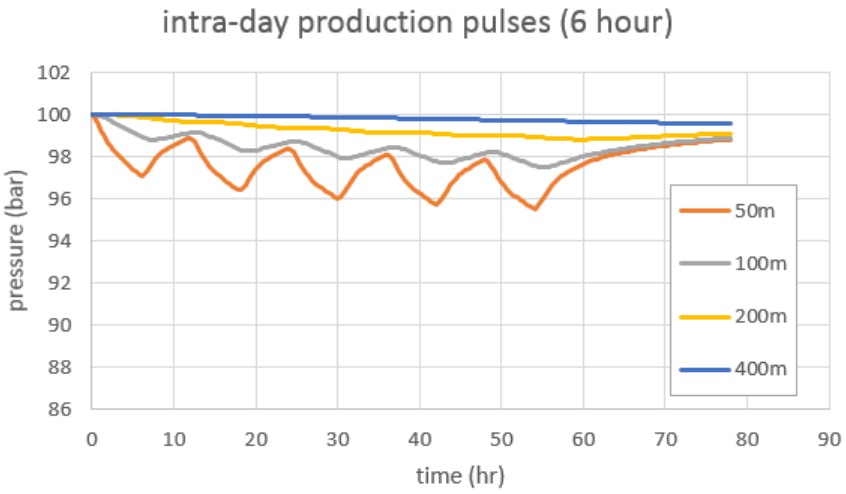


Figure 9-6: Numerical grid setup for testing the effect of a pressure pulse in observation wells at increasing distance (Voronoi grid)



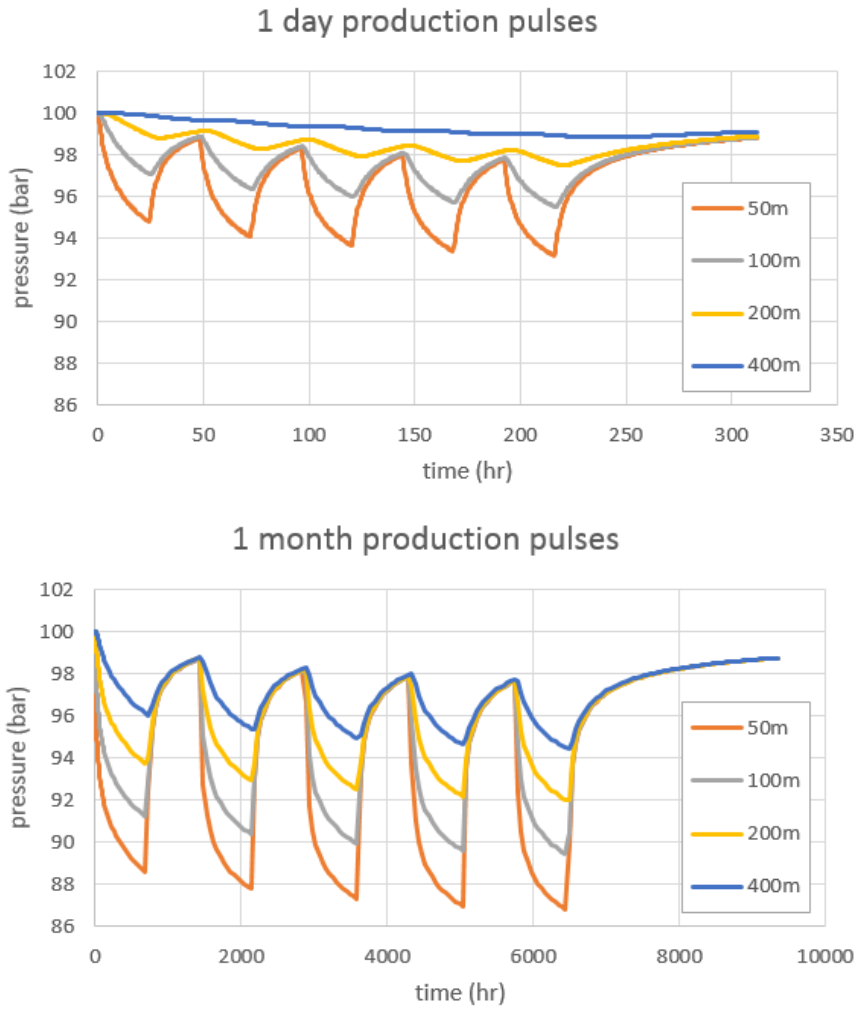


Figure 9-7: Dampening of a pressure transient in a porous medium with a highly compressible fluid at increasing distances from the source.

10 Recommendations

The conclusions are given in the Management Summary at the start of this document.

The optimisation needs to be reviewed regularly as part of the Meet en Regelprotocol, based on new data becoming available and improved understanding of the field. In a next update of the optimisation the study results should be updated with an improved seismological model which is expected result in a higher lateral resolution (from 5 to 3km).

Ideally the impact of pressure resolution should be tested in a similar structured approach to the compaction analysis, Reference [5]. This will allow for making informed choices for the optimisation to either use a potentially overdetermined model (overestimate the ability to resolve things), or to smooth too much and overlook potential optimisations. One can work out the predictive performance for models with various resolution (temporal/spatial). If it can be demonstrated that there is no massive deterioration in model quality, one can choose to use a high-resolution model. Such a choice would require confidence in the prior (which is needed to populate the high-resolution model).

A test plan will be prepared to monitor the gradual implementation of the optimised distribution of production. This will include test periods during which currently closed-in clusters are opened up in a controlled manner to test the seismic response of the area.

The toolset developed as part of this study allows for a structured approach to evaluate the model response with respect to the model controls (chapter 6). This methodology can be explored further in the context of the Meet en Regel Protocol.

The models as used in this analysis were calibrated up to the end of 2016. It is recommended to continue updating/calibrating the models as new data and advances in scientific understanding become available.

11 References

[1] "Technical Addendum to the Winningsplan Groningen 2016 - Production, Subsidence, Induced Earthquakes and Seismic Hazard and Risk Assessment in the Groningen Field, PART I – Summary and Production," NAM, Assen, 1/4/2016.

[2] J. Tomic, Harkstede 2A, Microseismic Monitoring Summary, NAM, 2017.

[3] "Methodology, Optimisation of the production distribution over the Groningen field to reduce seismicity," NAM, Assen, 29 May 2017.

[4] R. Paleja and S. Bierman, "Measuring changes in earthquake occurrence rates in Groningen," Shell, Amsterdam, October 2016.

[5] S. Bierman and D. Randell, Reservoir compaction estimation using intrinsic Gaussian Markov Random Field Models, Shell, July 2017.

[6] S. Bourne and S. Oates, "An activity rate model of induced seismicity within the Groningen field," 30 September 2014.

[7] S. Bourne and S. Oates, "Extreme threshold failures account for earthquakes induced by fluid extraction," Shell, 15/2/2017.

[8] H. Van Oeveren, P. Valvatne and L. Geurtsen, "Groningen Dynamic Model Updates 2017," NAM, Assen, 2017.

[9] S. Bourne and S. Oates, "An activity rate model of seismicity induced by reservoir compaction and fault reactivation in the Groningen gas field," Shell, 22 June 2015.

[10] "Study and Data Acquisition Plan Induced Seismicity in Groningen Update Post-Winningsplan 2016," NAM, Assen, 1/4/2016.

[11] G. Gao, "An Efficient Optimisation Work Flow for Field-Scale In-Situ Upgrading Developments," August 2015.

[12] G. Gao, "A Parallelized and Hybrid Data-Integration Algorithm for History Matching of Geologically Complex Reservoirs," December 2016.

[13] D. Wald, V. Quitoriano, T. Heaton and H. Kanamori, Relationships between peak ground acceleration, peak ground velocity, and modified Mercalli intensity in California, Earthquake Spectra 15, 557-564., 1999.

[14] D. Wald, K. Jaiswal, K. Marano, D. Bausch and M. Hearne, PAGER — Rapid assessment of an earthquake's impact., U.S. Geological Survey Fact Sheet 2010–3036, 2010.

[15] T. I. Allen, D. Wald, P. S. Earle, K. D. Marano, A. Hotovec, K. Lin and M. Hearne, An Atlas of ShakeMaps and population exposure catalog for earthquake loss modeling., Bull. Earthq. Eng., v. 7, DOI: 10.1007/s10518-009-9120-y., 2009.

[16] J. J. Bommer and et al, "V4 Ground-Motion Model (GMM) for Response Spectral Accelerations, Peak Ground Velocity, and Significant Durations in the Groningen Field," NAM, Assen, June 2017.

[17] L. Devroye, "Non-Uniform Random Variate Generation," Springer-Verlag, New-York, 1986.

[18] L. Breiman, "Random Forests," *Machine Learning,* pp. 5-32, 2001.

[19] "Groningen Meet- en Regelprotocol," NAM, Assen, 29/5/2017.

[20] Q. de Zeeuw, "CITHP to CIBHP conversion for the Groningen wells," NAM, 26 April 2016.

[21] G. Gao, J. Vink, F. Alpak and W. Mo, "An Efficient Optimisation Workflow for Field-Scale In-Situ Upgrading Developments," Society of Petroleum Engineers. doi:10.15530/urtec-2014-1885283., 2014.

[22] P. E. Gill, W. Murray and M. H. Wright, *Practical Optimisation,* London: Academic Press, 1981.

[23] C. Broyden, *Quasi-Newton methods and their application to function minimisation,* vol. 21, 1967, pp. 368-381.

[24] T. Mitchell, Machine Learning, McGraw Hill. ISBN 0-07-042807-7., 1997.

[25] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction, Springer. ISBN 978-0-387-84857-0, 2008.

[26] L. Breiman, "Random Forests," January 2001. [Online]. Available: https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf.

[27] A. Liaw, "The randomForest package for R - Breiman and Cutler's Random Forests for Classification and Regression," 7th October 2015. [Online]. Available: https://cran.r-project.org/web/packages/randomForest/randomForest.pdf.

[28] M. Kursa and W. Rudnicki, "Feature Selection with The Boruta Package," *Journal of Statistical Software,* vol. 36, no. 11, 2010.

[29] J. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics,,* vol. 29, pp. 1189-1232, 2001.

[30] K. Coppersmith, "Report on Mmax Expert Workshop, Mmax panel chairman Kevin Coppersmith," 2016.

Appendix A Cluster abbreviations

Table A-1: List of cluster abbreviations and locations. Locations are given in UTM coordinate system, Reference Amersfoort RD.

Name location	Abbreviation	X (m)	Y (m)
Amsweer	AMR	256264	591481
Bierum	BIR	254744	599407
Eemskanaal	EKL	241539	584421
De Eeker	EKR	259467	577265
Froombosch	FRB	248233	578952
Kooipolder	KPD	246529	580964
Leermens	LRM	250213	597054
Overschild	OVS	250484	590777
Oudeweg	OWG	256115	585671
De Paauwen	PAU	246052	588368
Ten Post	POS	245635	591366
Sappemeer	SAP	249530	575385
Schaapbulten	SCB	257329	588412
Siddeburen	SDB	253054	587461
Slochteren	SLO	246416	579285
Spitsbergen	SPI	252380	577234
Scheemderzwaag	SZW	257063	578196
Tjuchem	TJM	254927	588194
Tusschenklappen	TUS	254433	575176
‘t Zandt	ZND	247918	600637
Zuiderpolder	ZPD	261840	581024
Zuiderveen	ZVN	252967	579304

Appendix B Compressibility model

B.1 Data

The available data to establish a compressibility model are:

- Subsidence measurements (surface imprint of reservoir compaction)
  Leveling surveys, InSAR data, GPS stations
- In-situ compaction measurements
  Time lapse logging of marker bullets in observation wells
- Experimental measurements of rock compressibility on core samples

Subsidence measurements offer the only dataset of both temporal and spatial resolution, and dominates the rock compressibility in the HRA model. As a consequence, the lateral ability to resolve the rock compressibility is limited by the reservoir depth (Figure B-1).

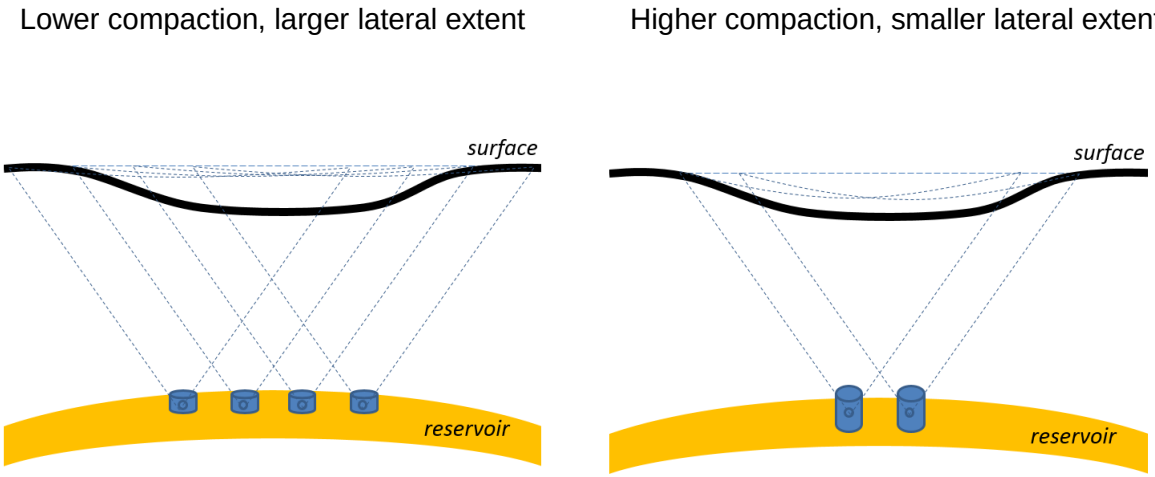


Figure B-1: Schematic illustration of filtering effect of overburden, from reservoir compaction to surface subsidence

B.2 Statistical evaluation of compressibility model

Reservoir compaction is thought to be one of the most important contributing factors to earthquakes in the Groningen field. Sound statistical methodology is required for compaction forecasting with realistic uncertainty estimates. The compressibility model (grid) that is used in the seismicity modelling was thoroughly evaluated for statistical trends, Reference [5]. Five compressibility models were cast in a Bayesian statistical framework for proper uncertainty evaluation:

- Reference model (constant reservoir compressibility in space and time)
- Winningsplan 2013
- Winningsplan 2016a
- Winningsplan 2016b
- Markov Random Field

The predictive performance of all models was compared for combinations of:

- Reservoir grid resolutions
- Time decay parameters
- Different priors for compressibilities (where applicable)

The reference model was found to do well overall, but is spatially biased. The other models at best do only a bit better. The best predictive performance was found for coarser reservoir grid resolutions (e.g. 3000m x 3000m or coarser). Porosity-compressibility priors do not make models perform better. There is no information in the surface displacement measurements to estimate compressibility or compaction at high grid resolutions, unless you have a strong belief in the porosity-compressibility relationship.

Recommendations for production optimisation study:

- Work with both the Reference model and IGMRF model at coarse resolution.
- Merge both models (future work).
- Possibly: use the Winningsplan 2016 models, but it would be good to estimate the strength of the prior (at the moment needs to be set 'by hand').

B.3 Compressibility grid used in Production Optimisation study

Based on the analysis, the assigned compressibility is the result of a geomechanical inversion from observed subsidence, using among other things the modelled reservoir pressure as input (Figure B-2). Essentially, we are now integrating the levelling data used for subsidence prediction into the dynamic reservoir model. The compressibility model is the same as used by the models used for the 2016 Hazard and Risk Assessment.

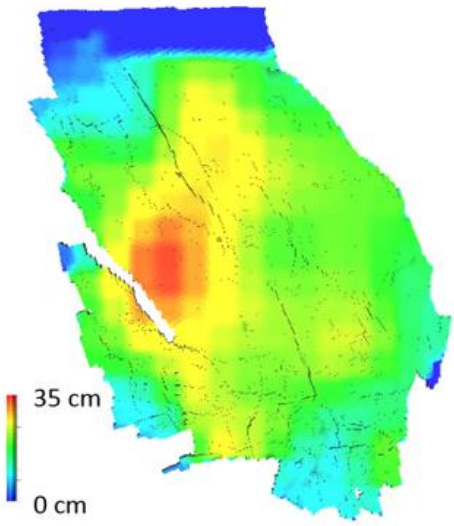


Figure B-2: Estimated compaction at 31st Dec 2016 using subsidence inversion. The compressibility values used are assuming a linear model in pressure.

Appendix C Optimisation

C.1 Optimisation algorithm

The mathematical optimisation was executed using the Simultaneous Perturbation and Multivariate Interpolation (SPMI) algorithm. SPMI is one of Shell's in-house developed optimisation methods [21]. SPMI is efficient for optimisation problems in which forward model evaluations are computationally expensive. The parallel nature of SPMI's implementation renders it particularly suitable for problems when analytic/adjoint gradient information is unavailable. When the analytical derivative of the objective function with respect to a variable is not available, as in this optimisation study, a finite difference approximation with auto-adaptive perturbation size method is used to estimate the derivative. These numerical gradients are computed in parallel at a given iteration, incurring a run-time cost (elapse time) approximately equivalent to a single forward model evaluation.
Without Multivariate Interpolation (MI) searching points, the SPMI algorithm is almost the same as a direct pattern search - a derivative-free method. The auto-adaptive perturbation size updating procedure makes the direct search more efficient because it may generate a more appropriate pattern size. If the objective function exhibits some degree of smoothness or trends, the MI searching points guided by a local quadratic approximation of the true objective function should improve the efficiency of a local search significantly. Therefore, the SPMI optimisation tool can claim the advantages of both direct pattern search (suitable for non-smooth functions) and model-based local search (more efficient for a smooth function).

To summarize, the SPMI algorithm is a hybrid method combining a trust region quasi-Newton method (see Reference [22] for its definition) and direct pattern search method. It exhibits the following features:

- Simultaneous Perturbation – generates 2n (n is the number of optimisation variables) simultaneous perturbation points of the base case (current iteration), 2-sided perturbations are performed per parameter using an auto-adaptive perturbation size (expand if the objective function improves at both perturbation points or shrunk if the objective function does not improve at both perturbation points).
- Multivariate Interpolation – constructs an approximate quadratic model with application of a quasi-Newton technique (using Symmetric Rank one, for Hessian approximations - [23]), and subsequently generates m searching points, to search within trust-regions. MI searching points (generated using three strategies: ball-shaped trust region, box-shaped trust region, and a line search) are guided by that local quadratic approximation of the objective function.
- Random search points – generates r random searching points to provide opportunity of escaping from a local optimum (by randomly exploring the solution space).
- Parallelization – submits all $2n+3m+r$ cases simultaneously.

The Simultaneous Perturbations and Multivariate Interpolation (SPMI) method has been used to optimise the initial control strategies.

During each SPMI optimisation the random seed of the HRA is fixed to 1. Changing the random seed during SPMI iterations introduces noise into the gradient calculation of the SPMI method and results in a slower convergence. However, to see the impact of the random seed on SPMI results, we repeat the SPMI optimisation with two different random seeds (value set to 1 and 2). See Figure C-1 for comparison of the resulting strategies, which are fairly similar.

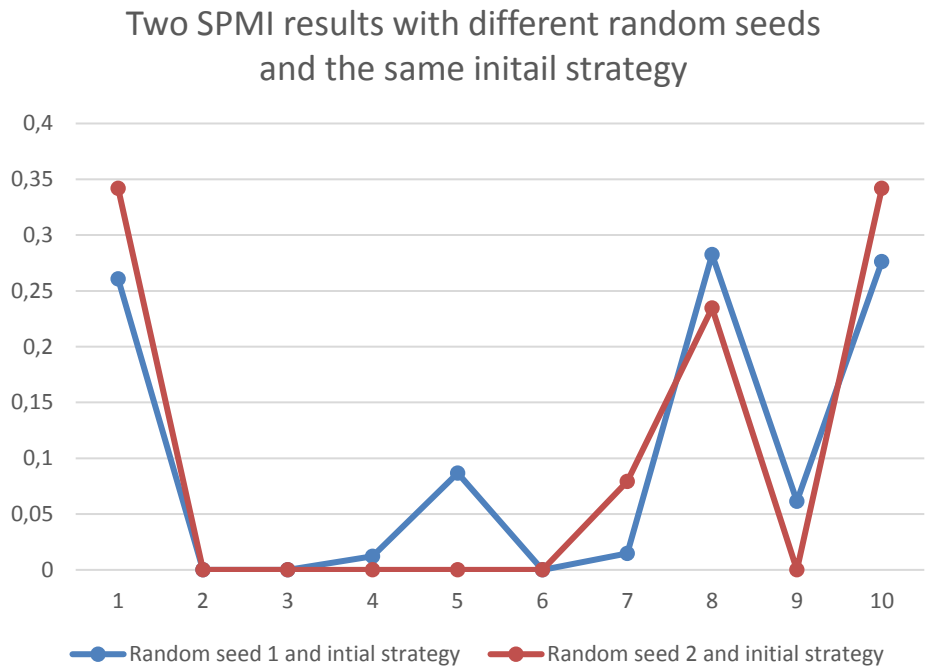


Figure C-1 Comparison of two SPMI results generated with two different random seeds and BP17 production strategy. On x-axis: control number, on y-axis: optimal control value.

As an example, Table 2 shows the results of five SPMI optimisations (with respect to pwPGV) where the best results of the random sampling procedure (described in section 0) were used as initial strategies. In general, SPMI finds further improvement above the HRA noise level but does not go far from the solutions explored during the random sampling. Similar low values of the metrics can be achieved by significantly different control strategies.

Table 2 Results of 5 SPMI optimisations with different initial strategies

case	Mean PGV	Max PGA	Max PGV	Tremor	Avg prod	Cluster_1	Cluster_2	Cluster_3	Cluster_4	Cluster_5	Cluster_6	Cluster_7	Cluster_8	Cluster_9	Cluster_10
1: init	0.052883	0.264434	0.166381	81.64672	21.586	0.483843	0.042577	0.003065	0.018613	0.092863	0.012811	0.040025	0.243869	0.060716	0.001617
1: opt	**0.04941**	**0.243886**	**0.157156**	**76.25336**	**21.43**	**4.68E-01**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**8.98E-02**	**1.24E-02**	**3.87E-02**	**3.32E-01**	**5.87E-02**	**0.00E+00**
2: init	0.052871	0.266102	0.171391	82.42058	2.15E+01	0.380784	0.100552	0.002881	0.010548	0.033115	0.064309	0.017066	0.316425	0.046843	0.027476
2: opt	**0.050144**	**0.257367**	**0.162873**	**81.31988**	**2.15E+01**	**0.397226**	**0.104893**	**0**	**0.011004**	**0.034545**	**0.067086**	**0.017803**	**0.330096**	**0**	**0.037347**
3: init	0.052405	0.251667	0.163113	84.489	2.14E+01	0.011228	0.005722	0.018836	0.03725	0.224157	0.164357	0.022054	0.340053	0.009404	0.166938
3: opt	**0.051117**	**0.251416**	**0.157852**	**84.10494**	**2.15E+01**	**0.016591**	**0**	**0.016773**	**0.040647**	**0.219221**	**0.167178**	**0.024777**	**0.334356**	**0.011142**	**0.169316**
4: init	0.052047	0.258878	0.169119	82.7089	2.16E+01	0.496692	0.022515	0.01267	0.023344	0.159091	0.034765	0.004989	0.194278	0.012256	0.039401
4: opt	**0.049499**	**0.254125**	**0.16304**	**78.73048**	**2.16E+01**	**0.420316**	**0.023727**	**0**	**0**	**0.168608**	**0.036189**	**0.004962**	**0.270292**	**0.035325**	**0.040581**
5: init	0.052619	0.262312	0.167934	85.19996	2.15E+01	0.517315	0.034522	0.015284	0.018371	0.082645	0.031135	0.023634	0.161609	0.043676	0.071808
5: opt	**0.04901**	**0.245899**	**0.162361**	**80.19968**	**2.16E+01**	**0.360611**	**0.034073**	**0**	**0**	**0.167122**	**0.10815**	**0**	**0.261916**	**0**	**0.068127**

Figure C-2 depicts two SPMI results generated with two different random seeds for BP17 production strategy together with five SPMI results generated with random seed equal 1 and different initial

production strategies explored during the random sampling (and captured in Table 2). These control strategies result in very similar metrics values. Figure C-2 gives a rough idea about the null space of the solution. See section 0 for more details.

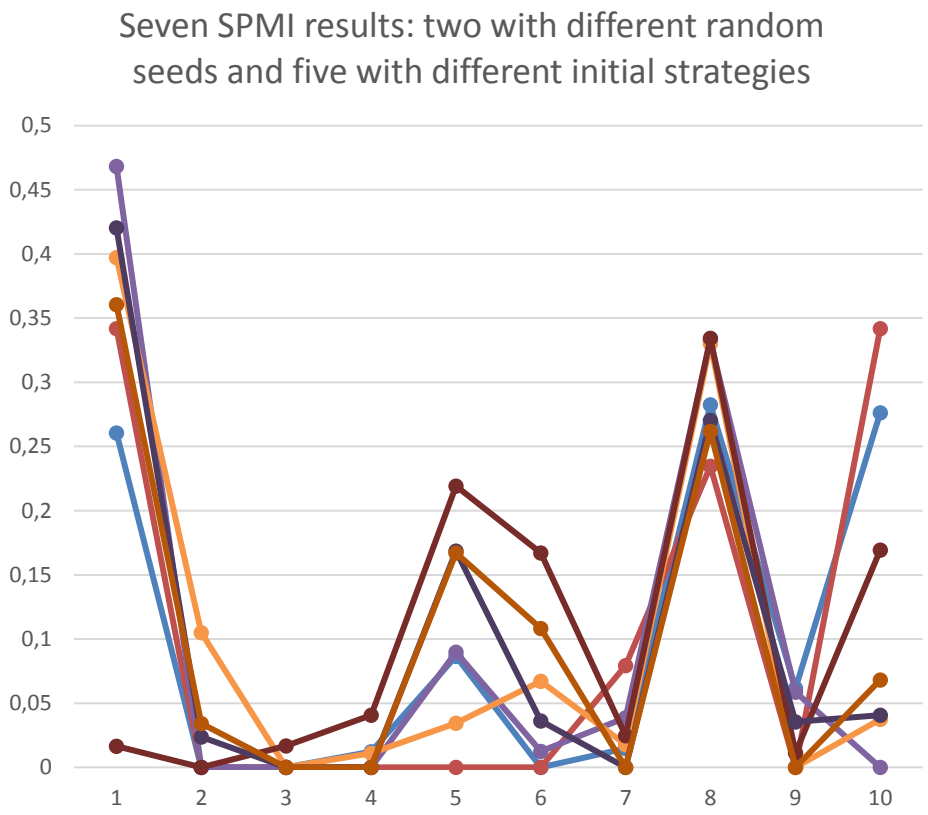


Figure C-2 Comparison of different SPMI results resulting from changing random seed or changing starting production strategy. On x-axis: control number, on y-axis: optimal control value

C.2 Normalization of the objective function

For the purpose of optimiser efficiency, the objective metric is normalized. This also allows for combining two different metrics in one objective function, e.g. Hybrid metric combining pwPGV with Event Count.

For each of the branches in the hazard assessment model, the objective function is scaled using a reference value which is associated with on the initial production distribution control settings. To illustrate, when consider the pwPGV metric, the objective function becomes:

$$J_{seismicity} = \frac{pwPGV}{ref\ pwPGV}$$

The pwPGV will fluctuate over the various runs throughout the optimisation as a result of the changes in production distribution control settings, while ref pwPGV remains constant. It works as a scaling factor, and brings the objective function to 1 for the initial control strategy.

To bring the objective function to the same order of magnitude as the total number of events, they can be scaled by a factor of 100.

C.3 Penalty term

The underlying seismicity model relates depletion to compaction, stress build-up, earthquakes and ground movement. The seismicity metrics reduce with the decrease in total field production rate. Consequently, a mathematical optimiser will steer towards zero production when left unconstrained.

To prevent the optimisation from going into the solutions that tend to decrease the field production rate, a penalty term was added to the objective function. This penalty term kicks in if the aspired total field production rate is not met. In Figure C-3 it can be seen how much the penalty term increases if production deviates from the 21.6 Bcm constraint. Penalty terms of the following form have been used:

1) $J_{penalty} = a(21.6 - r_{actual}) + b\left(\frac{21.6 - r_{actual}}{c}\right)^d$,
2) $J_{penalty} = b\left(MAX(\frac{21.6 - r_{actual}}{c} - e, 0)\right)^d$,

where r_{actual} is an achieved averaged (over the forecast period) yearly production rate, a, b, c, d and e are the penalty term parameters. Because of the modeling setup r_{actual} is always less or equal to the aspired rate. Parameter e is related to deviation from the target field production rate. The value of 0.005 corresponds to 0.1 Bcm per year deviation from the target field production rate.

The following parameter sets have been used:

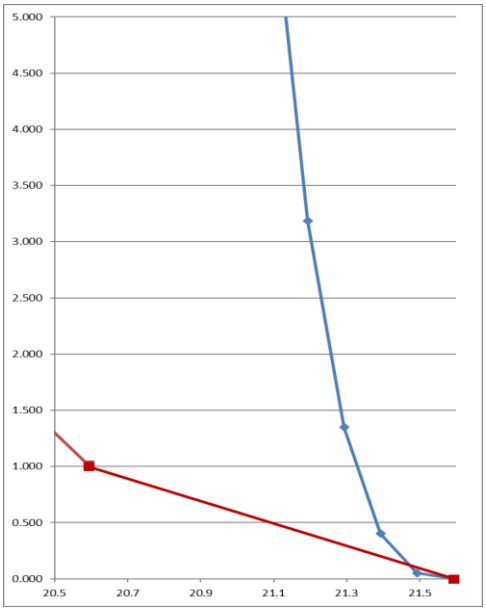
1) $a = 0.02$, $b = 500000$, $c = 3$ and $d = 3$,
2) $a = 0$, $b = 50000$, $c = 21.6$ and $d = 2$, $e = 0.005$.

To identify a useful range of these parameters several simulations were run with reduced yearly production rates. For a small drop in the production rate, the penalty term is flat or deviates insignificantly as compared to the expected reduction in the seismic metric due to lower field production rate, but it steeply increases when production deviates far from 21.6 Bcm per year (see Figure C-3, in which all clusters have been reduced proportionally to achieve the investigated rate).

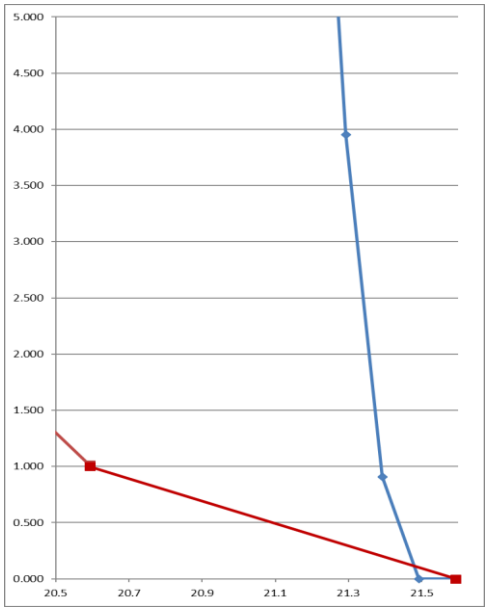
To summarize, the structure of the penalty term is chosen to

1) allow the optimisation to optimise only on seismicity
2) be strictly increasing for increasing deviation from the target field production rate
3) become dominant quickly above a specific deviation.

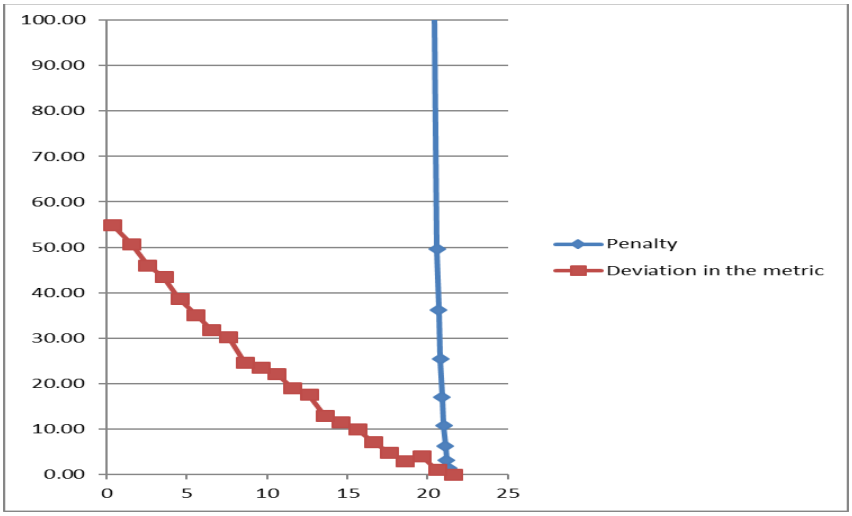
This parameterization worked well for the executed studies. However, it might require some adjustment if the production strategies change.



(a) Penalty term (1)



(b) Penalty term (2)



c) Observed decrease in metric with deviation from 21.6 Bcm per year and penalty term

Figure C-3 On x-axis: production rate per year, on y-axis: in red observed decrease in metric with deviation from 21.6 Bcm per year, in blue: penalty term.

To give an example, the objective function for pwPGV and the second type penalty term has the following form

$$J(c) = J_{seismic}(c) + J_{penalty}(c)$$

$$= 100 \cdot \frac{pwPGV(c)}{ref\ pwPGV} + 50000\left(MAX(\frac{21.6 - r_{actual}(c)}{21.6} - 0.005, 0)\right)^2$$

where c in this case is a vector in $\mathbb{R}^{10}$ consisting of fractions of the total production rate assigned to each of 10 groups of production clusters, i.e. $0 \leq c_i \leq 1$ and $\sum_{i=1}^{10} c_i = 1$.

Appendix D Optimisation details

D.1 Event count metric

Table D-1: Seismic metrics for optimised distribution based on tremor count objective. Percentage improvement is w.r.t. BP17 reference case.

Branch	pwPGV	maxPGA	maxPGV	# EQ	pwPGV	maxPGA	maxPGV	# EQ
1	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
2	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
3	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
4	0.045	0.226	0.134	95	-13%	-10%	-6%	-17%
5	0.045	0.226	0.133	95	-13%	-10%	-7%	-17%
6	0.045	0.226	0.133	95	-13%	-10%	-7%	-17%
7	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
8	**0.045**	**0.225**	**0.133**	**95**	**-14%**	**-11%**	**-7%**	**-17%**
9	0.045	0.224	0.132	95	-13%	-11%	-7%	-17%
10	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
11	0.045	0.227	0.133	95	-13%	-10%	-7%	-17%
12	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
13	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
14	0.045	0.227	0.134	95	-13%	-10%	-6%	-17%
15	0.045	0.225	0.133	95	-13%	-11%	-7%	-17%
16	0.045	0.226	0.133	95	-13%	-10%	-7%	-17%
17	0.044	0.227	0.137	95	-15%	-10%	-4%	-17%
18	0.044	0.228	0.136	95	-15%	-10%	-5%	-17%
19	0.044	0.228	0.137	96	-14%	-10%	-4%	-17%
20	0.044	0.227	0.137	95	-15%	-10%	-4%	-17%
21	0.044	0.227	0.137	95	-15%	-10%	-4%	-17%
22	0.044	0.229	0.137	95	-14%	-9%	-4%	-17%
23	0.044	0.227	0.137	95	-14%	-10%	-4%	-17%
24	0.044	0.227	0.137	95	-14%	-10%	-4%	-17%

Table D-2: Optimised normalized production splits for different logic tree branches for tremor count objective.

Branch	East	OPP	EKL	LRM	BIR	ZND	S-West	S-East	SDB	S-Cnt
1	0.50	-	-	-	-	-	-	0.29	-	0.21
2	0.50	-	-	-	-	-	-	0.29	-	0.21
3	0.50	-	-	-	-	-	-	0.29	-	0.21
4	0.50	-	-	-	-	-	-	0.29	-	0.21
5	0.50	-	-	-	-	-	-	0.29	-	0.21
6	0.50	-	-	-	-	-	-	0.29	-	0.21
7	0.50	-	-	-	-	-	-	0.29	-	0.21
8	**0.50**	**-**	**-**	**-**	**-**	**-**	**-**	**0.29**	**-**	**0.21**
9	0.51	-	-	-	-	-	-	0.29	-	0.20
10	0.51	-	-	-	-	-	-	0.29	-	0.20
11	0.51	-	-	-	-	-	-	0.29	-	0.20
12	0.51	-	-	-	-	-	-	0.29	-	0.20
13	0.51	-	-	-	-	-	-	0.29	-	0.20
14	0.51	-	-	-	-	-	-	0.29	-	0.20
15	0.51	-	-	-	-	-	-	0.29	-	0.20
16	0.51	-	-	-	-	-	-	0.29	-	0.20
17	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07
18	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07
19	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07
20	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07
21	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07
22	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07
23	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07
24	0.50	-	-	-	0.10	-	-	0.31	0.03	0.07

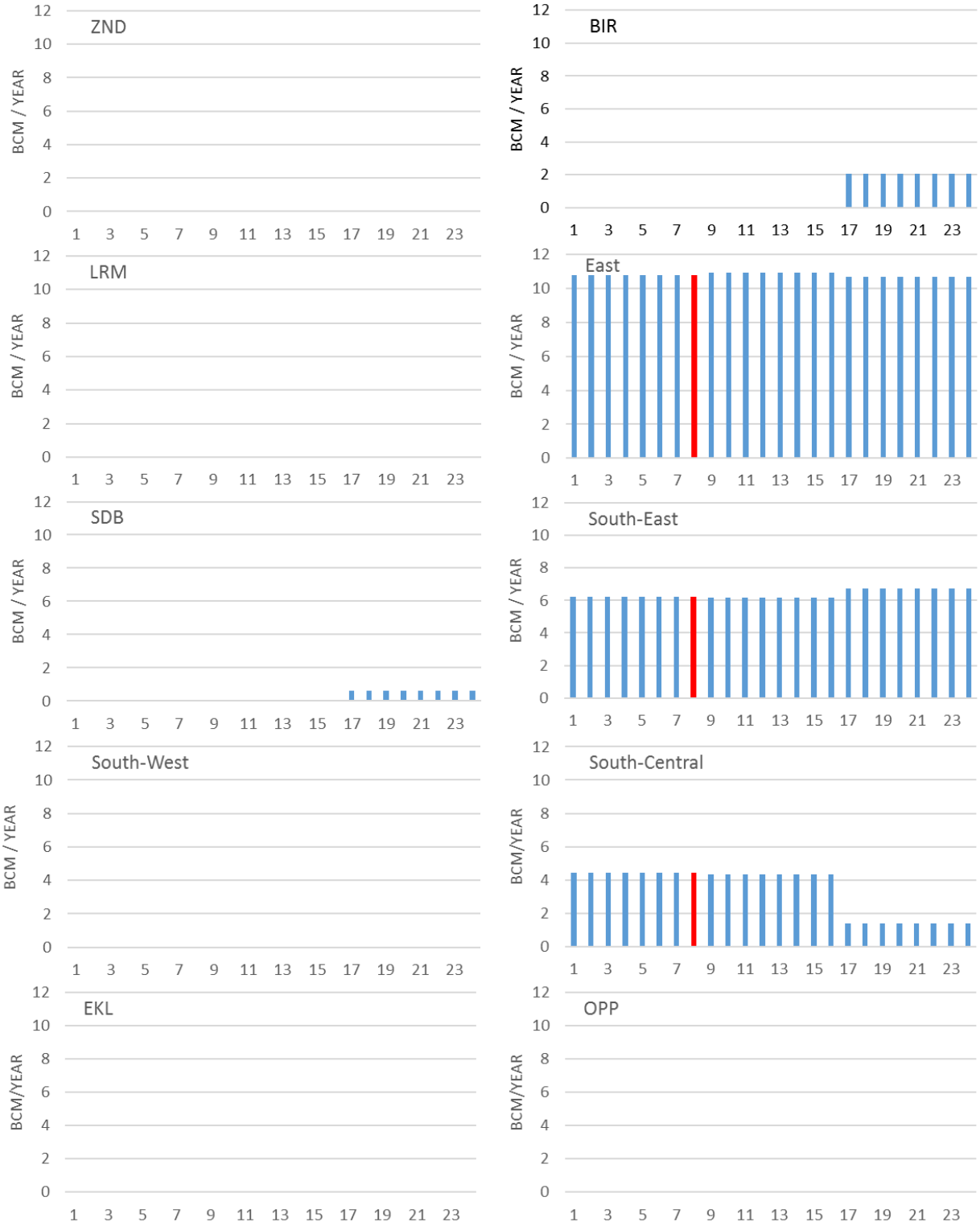


Figure D-1: Optimised production splits for different logic tree branches for tremor count metric. Branch 8, indicated in red, gave the optimal solution when evaluated in terms of the logic tree mean.

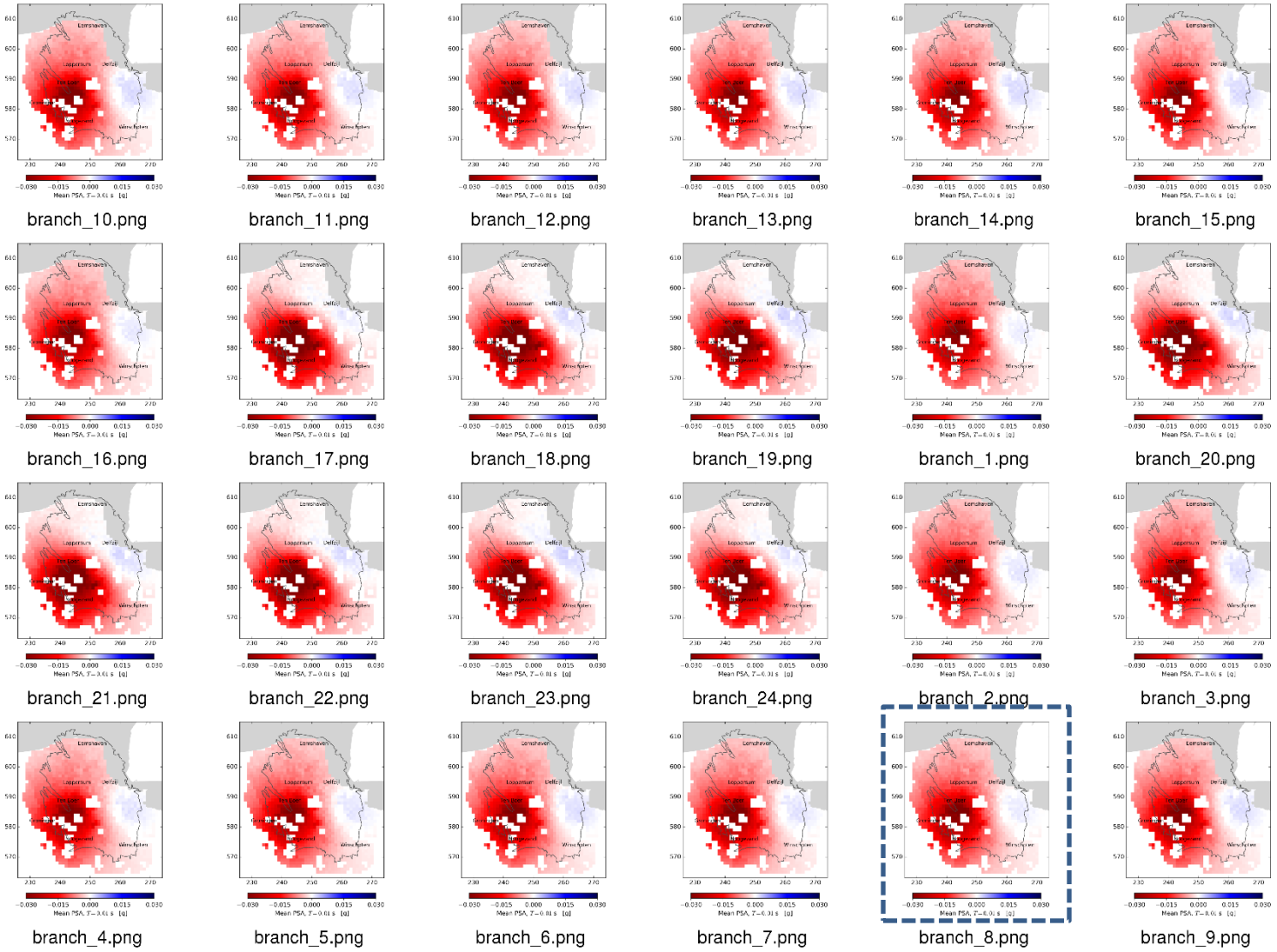


Figure D-2: Mean PGA hazard difference maps (scaled in [-0.03,0.03]) for Event-count driven optimisation, at average 0.2% annual chance of exceedance (1 in 475 years) from 1/1/2018 to 31/12/2022 (optimal solution indicated)

D.2 Max PGA

Table D-3: Seismic metrics for optimised distribution based on max PGA objective. Percentage improvement is w.r.t. BP17 reference case.

Branch	pwPGV	maxPGA	maxPGV	# EQ	pwPGV	maxPGA	maxPGV	# EQ
1	0.046	0.226	0.133	97	-11%	-10%	-7%	-16%
2	0.045	0.230	0.138	98	-12%	-9%	-3%	-14%
3	0.048	0.233	0.135	104	-6%	-8%	-5%	-9%
4	0.047	0.230	0.134	101	-8%	-9%	-6%	-12%
5	0.046	0.229	0.135	100	-10%	-9%	-5%	-13%
6	0.046	0.228	0.133	98	-11%	-10%	-7%	-15%
7	0.046	0.226	0.133	97	-11%	-10%	-7%	-16%
8	0.046	0.228	0.135	98	-11%	-10%	-5%	-15%
9	0.046	0.228	0.134	99	-10%	-9%	-6%	-14%
10	0.045	0.230	0.139	99	-12%	-9%	-3%	-14%
11	0.046	0.228	0.136	99	-12%	-10%	-5%	-14%
12	0.046	0.226	0.134	98	-11%	-10%	-6%	-15%
13	0.045	0.229	0.138	98	-13%	-9%	-4%	-15%
14	0.047	0.230	0.137	100	-10%	-9%	-5%	-13%
15	0.044	0.228	0.140	97	-15%	-9%	-2%	-16%
16	**0.045**	**0.226**	**0.135**	**96**	**-13%**	**-10%**	**-5%**	**-16%**
17	0.046	0.229	0.135	99	-10%	-9%	-6%	-14%
18	0.045	0.230	0.136	97	-12%	-9%	-5%	-15%
19	0.046	0.228	0.134	99	-10%	-10%	-7%	-14%
20	0.046	0.234	0.139	102	-10%	-7%	-3%	-11%
21	0.046	0.231	0.138	100	-12%	-8%	-4%	-13%
22	0.046	0.230	0.140	100	-12%	-9%	-2%	-13%
23	0.045	0.231	0.140	100	-12%	-8%	-2%	-13%
24	0.044	0.230	0.140	98	-14%	-9%	-2%	-15%

Table D-4: Optimised normalized production splits for different logic tree branches for max PGA objective.

Branch	East	OPP	EKL	LRM	BIR	ZND	S-West	S-East	SDB	S-Cnt
1	0.44	-	-	-	-	-	-	0.25	-	0.31
2	0.41	0.01	-	-	0.11	-	0.05	0.31	0.04	0.07
3	0.39	-	-	-	-	-	0.13	0.17	0.04	0.27
4	0.24	-	0.01	-	0.10	0.01	0.04	0.29	0.03	0.28
5	0.32	-	-	-	-	0.07	-	0.22	0.04	0.34
6	0.37	-	-	-	0.01	-	0.01	0.27	0.03	0.30
7	0.39	-	-	-	-	-	-	0.28	0.04	0.28
8	0.33	-	-	-	0.13	-	0.05	0.30	-	0.18
9	0.27	-	-	0.01	0.06	-	0.06	0.31	0.05	0.23
10	0.42	-	-	-	0.18	-	0.08	0.31	0.02	-
11	0.37	0.01	-	-	0.10	0.03	0.05	0.28	-	0.17
12	0.29	-	-	-	0.12	-	0.02	0.29	-	0.29
13	0.39	-	-	-	0.15	-	0.05	0.30	0.03	0.08
14	0.31	-	0.01	-	0.11	0.01	0.04	0.30	0.02	0.20
15	0.48	-	-	-	0.23	0.01	-	0.27	-	0.01
16	**0.33**	**0.01**	**-**	**-**	**0.15**	**-**	**-**	**0.31**	**-**	**0.21**
17	0.40	-	-	-	0.05	-	0.08	0.29	0.04	0.13
18	0.49	-	-	-	-	-	0.04	0.29	0.06	0.12
19	0.40	-	-	-	0.02	-	0.09	0.31	0.04	0.15
20	0.49	0.02	0.03	0.01	0.06	0.01	-	0.33	0.05	0.02
21	0.44	0.05	-	-	0.06	0.01	0.07	0.31	-	0.06
22	0.26	-	-	-	0.05	0.20	0.06	0.32	0.03	0.09
23	0.30	-	-	-	0.07	0.19	0.06	0.30	-	0.08
24	0.53	-	-	-	-	0.13	-	0.24	-	0.10

Figure D-3: Optimised production splits for different logic tree branches for max PGA metric. Branch 16, indicated in red, gave the optimal solution when evaluated in terms of the logic tree mean.

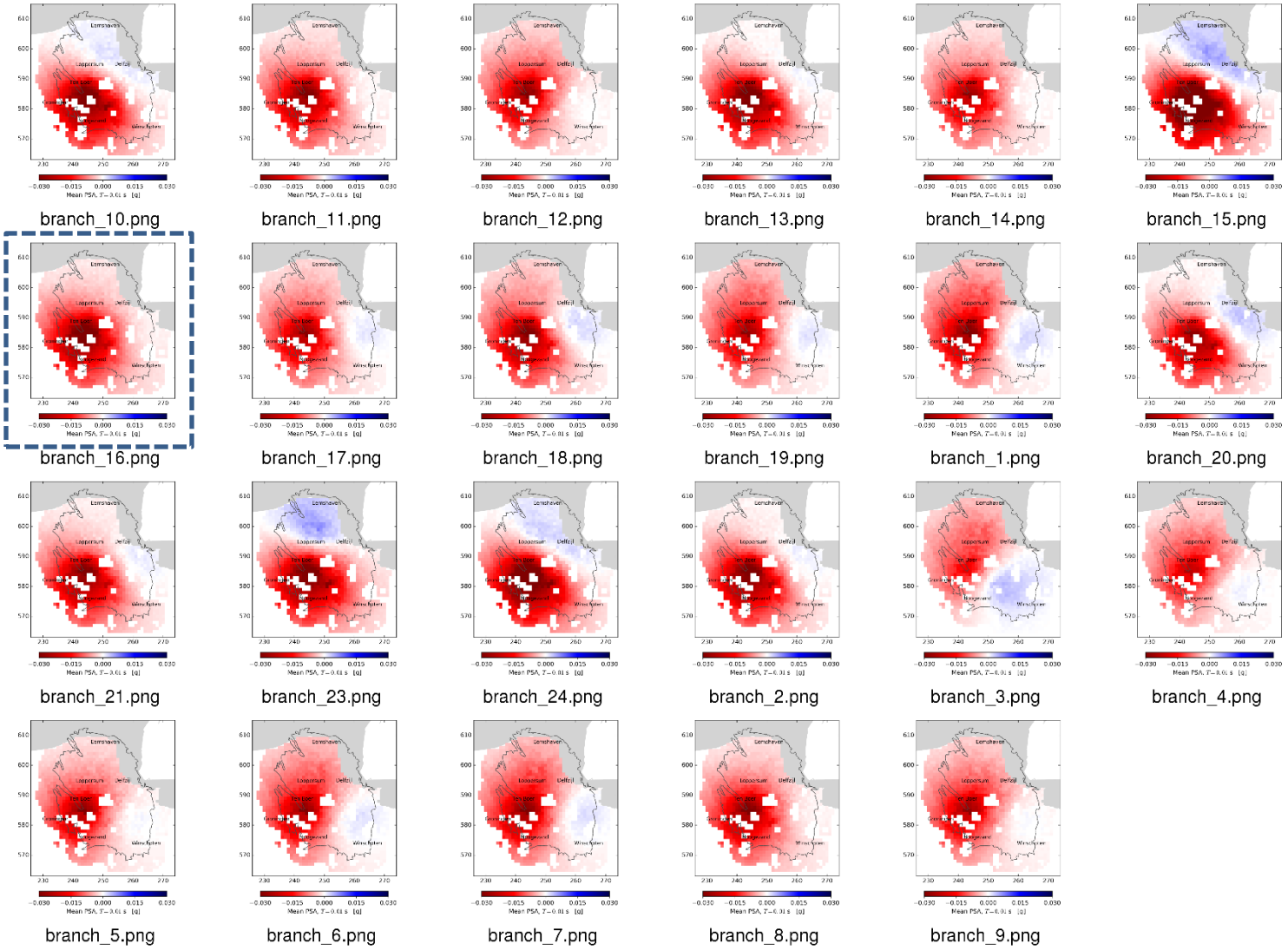


Figure D-4: Mean PGA hazard difference maps (scaled in [-0.03,0.03]) for Maximal-PGA driven optimisation, at average 0.2% annual chance of exceedance (1 in 475 years) from 1/1/2018 to 31/12/2022 (optimal solution indicated)

D.3 Max PGV

Table D-5: Seismic metrics for optimised distribution based on max PGV objective. Percentage improvement is w.r.t. BP17 reference case.

Branch	pwPGV	maxPGA	maxPGV	# EQ	pwPGV	maxPGA	maxPGV	# EQ
1	0.049	0.238	0.136	106	-4%	-6%	-5%	-8%
2	0.048	0.230	0.134	102	-8%	-9%	-6%	-11%
3	0.049	0.237	0.133	103	-5%	-6%	-7%	-10%
4	0.047	0.230	0.135	101	-8%	-9%	-6%	-12%
5	0.047	0.229	0.134	102	-8%	-9%	-6%	-12%
6	0.048	0.233	0.136	103	-7%	-8%	-5%	-11%
7	0.050	0.239	0.138	109	-3%	-5%	-3%	-5%
8	0.048	0.232	0.135	103	-7%	-8%	-6%	-11%
9	0.048	0.231	0.135	102	-8%	-8%	-6%	-11%
10	0.047	0.227	0.133	99	-10%	-10%	-7%	-14%
11	**0.047**	**0.228**	**0.132**	**99**	**-9%**	**-10%**	**-7%**	**-14%**
12	0.049	0.234	0.133	104	-6%	-7%	-7%	-10%
13	0.045	0.227	0.134	97	-12%	-10%	-7%	-16%
14	0.046	0.229	0.135	100	-10%	-9%	-6%	-13%
15	0.048	0.231	0.135	103	-7%	-8%	-6%	-11%
16	0.048	0.234	0.137	103	-7%	-7%	-4%	-10%
17	0.050	0.238	0.137	106	-4%	-6%	-4%	-7%
18	0.047	0.229	0.135	100	-9%	-9%	-5%	-13%
19	0.047	0.230	0.136	101	-9%	-9%	-5%	-12%
20	0.048	0.231	0.135	101	-8%	-8%	-6%	-12%
21	0.048	0.231	0.134	102	-7%	-8%	-6%	-11%
22	0.046	0.226	0.133	97	-11%	-10%	-7%	-15%
23	0.047	0.229	0.133	100	-9%	-9%	-7%	-13%
24	0.046	0.228	0.134	99	-10%	-10%	-6%	-14%

Table D-6: Seismic metrics for optimised distribution based on max PGV objective.

Branch	East	OPP	EKL	LRM	BIR	ZND	S-West	S-East	SDB	S-Cnt
1	0.21	-	-	0.08	-	-	0.19	0.27	-	0.25
2	0.17	-	-	-	0.06	-	0.09	0.29	0.12	0.27
3	0.29	-	-	-	-	-	0.19	0.27	-	0.25
4	0.19	-	-	-	0.05	-	0.07	0.29	0.13	0.26
5	0.20	-	-	-	0.07	0.02	0.09	0.28	0.07	0.27
6	0.17	0.02	-	-	0.06	-	0.07	0.28	0.11	0.29
7	0.17	0.02	0.03	0.01	0.04	0.01	0.07	0.28	0.13	0.25
8	0.18	-	-	-	0.06	-	0.12	0.30	0.12	0.23
9	0.29	-	-	-	-	0.01	0.19	0.26	-	0.25
10	0.37	-	-	-	-	-	0.02	0.24	0.05	0.32
11	**0.35**	**-**	**-**	**-**	**-**	**-**	**0.09**	**0.29**	**-**	**0.27**
12	0.22	-	-	-	0.03	-	0.17	0.28	0.05	0.25
13	0.43	0.02	-	-	-	-	-	0.28	-	0.27
14	0.33	-	-	-	0.04	-	-	0.23	0.08	0.31
15	0.22	0.04	-	-	0.02	-	0.20	0.29	0.04	0.19
16	0.32	-	0.02	-	-	-	-	0.24	0.10	0.32
17	0.18	0.02	0.03	0.01	0.10	0.01	0.07	0.30	0.01	0.27
18	0.19	-	-	-	0.13	-	0.07	0.29	0.04	0.28
19	0.14	-	-	0.02	0.14	-	0.07	0.30	0.05	0.27
20	0.32	-	0.02	-	0.05	-	-	0.28	0.04	0.28
21	0.18	0.01	-	-	0.10	-	0.11	0.29	0.04	0.26
22	0.39	-	0.01	-	0.02	-	-	0.29	0.02	0.26
23	0.27	-	-	-	0.05	-	0.06	0.28	0.04	0.30
24	0.34	-	0.01	-	0.03	-	0.04	0.29	0.03	0.26

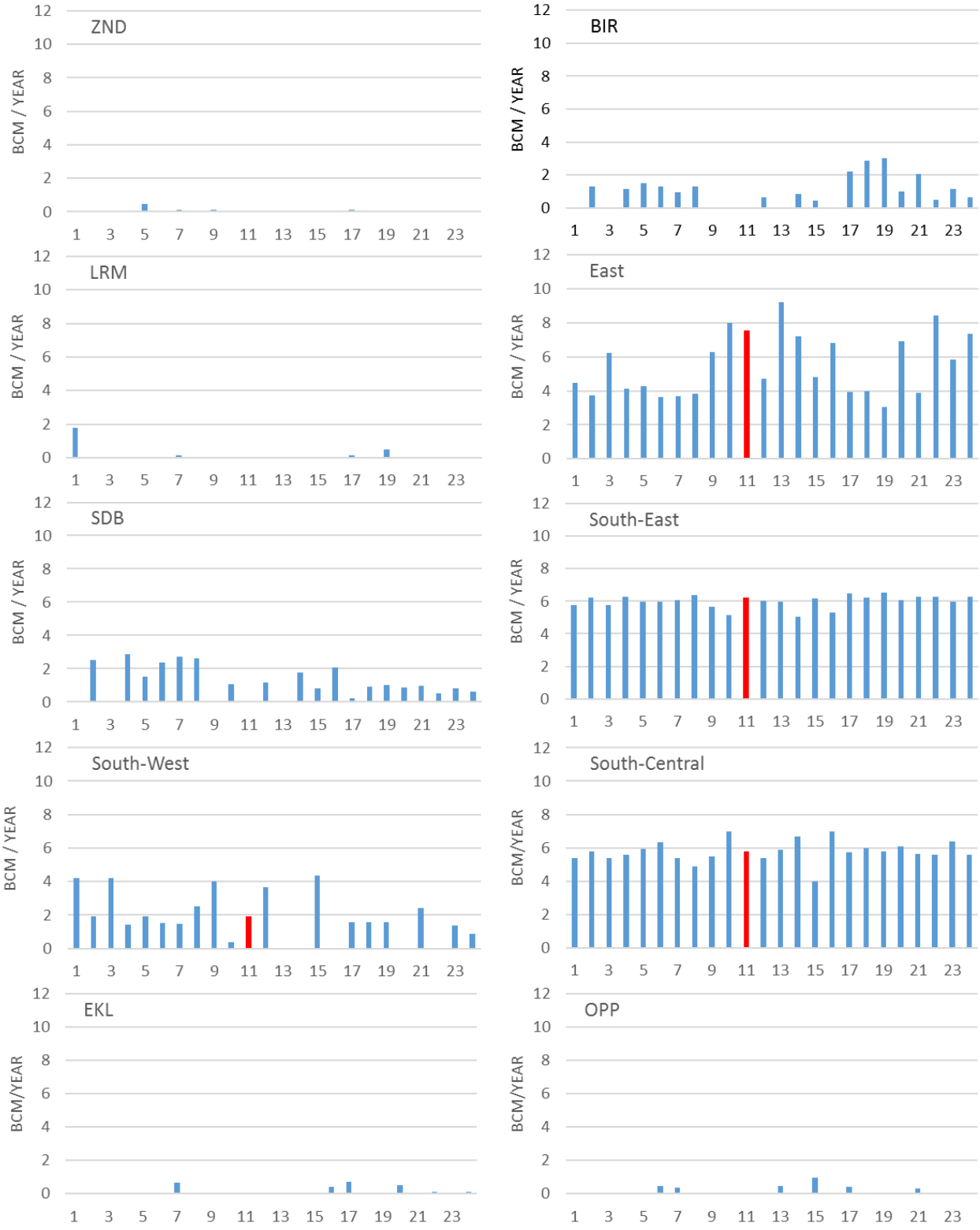


Figure D-5: Optimised production splits for different logic tree branches for max PGV metric. Branch 11, indicated in red, gave the optimal solution when evaluated in terms of the logic tree mean.

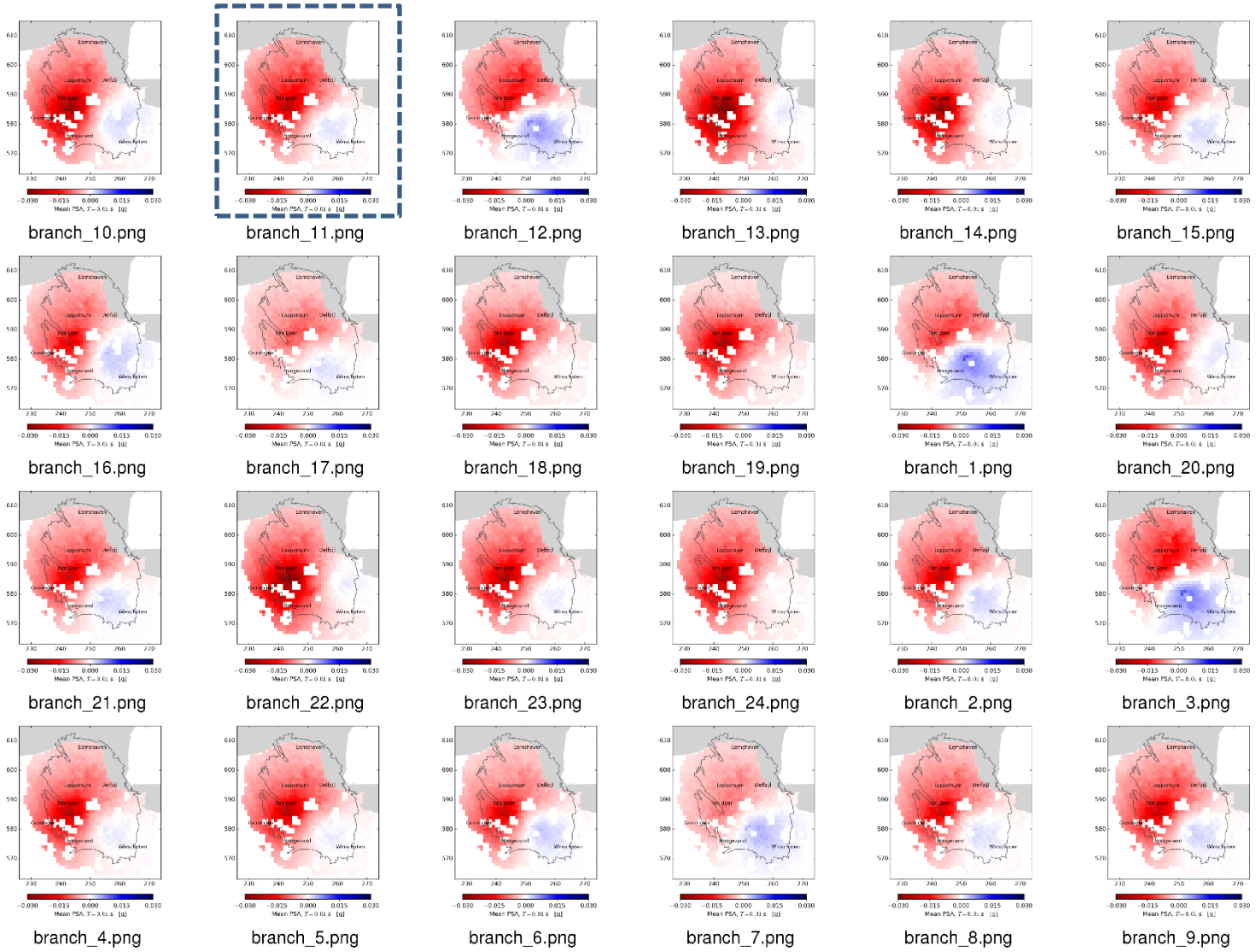


Figure D-6: Mean PGA hazard difference maps (scaled [-0.03,0.03]) for Maximal PGV driven optimisation, at average 0.2% annual chance of exceedance (1 in 475 years) from 1/1/2018 to 31/12/2022 (optimal solution indicated)

D.4 Population weighted PGV

Table D-7: Seismic metrics for optimised distribution based on population weighted PGV objective. Percentage improvement is w.r.t. BP17 reference case.

Branch	pwPGV	maxPGA	maxPGV	# EQ	pwPGV	maxPGA	maxPGV	# EQ
1	0.044	0.231	0.143	99	-15%	-8%	0%	-14%
2	0.045	0.235	0.144	101	-13%	-7%	0%	-12%
3	0.044	0.229	0.140	98	-14%	-9%	-2%	-15%
4	0.045	0.231	0.141	100	-12%	-8%	-2%	-13%
5	0.045	0.229	0.139	100	-13%	-9%	-3%	-13%
6	0.044	0.230	0.142	100	-14%	-9%	0%	-13%
7	**0.044**	**0.230**	**0.143**	**99**	**-15%**	**-9%**	**0%**	**-14%**
8	0.044	0.230	0.141	99	-15%	-9%	-1%	-14%
9	0.046	0.240	0.150	107	-11%	-5%	5%	-7%
10	0.046	0.242	0.153	109	-11%	-4%	7%	-5%
11	0.045	0.233	0.142	101	-13%	-8%	-1%	-12%
12	0.046	0.236	0.146	106	-10%	-6%	2%	-8%
13	0.045	0.233	0.143	101	-13%	-8%	0%	-12%
14	0.046	0.240	0.149	108	-10%	-5%	4%	-6%
15	0.045	0.235	0.148	105	-12%	-7%	3%	-9%
16	0.046	0.237	0.147	106	-10%	-6%	3%	-8%
17	0.046	0.236	0.146	106	-11%	-7%	2%	-8%
18	0.047	0.240	0.149	109	-9%	-5%	4%	-5%
19	0.047	0.242	0.151	110	-9%	-4%	6%	-4%
20	0.046	0.238	0.147	108	-10%	-6%	3%	-6%
21	0.047	0.242	0.149	110	-9%	-4%	5%	-4%
22	0.045	0.232	0.143	102	-13%	-8%	0%	-12%
23	0.048	0.248	0.154	115	-8%	-2%	8%	0%
24	0.048	0.250	0.155	114	-7%	-1%	9%	-1%

Table D-8: Optimised normalized production splits for different logic tree branches for population weighted PGV objective.

Branch	East	OPP	EKL	LRM	BIR	ZND	S-West	S-East	SDB	S-Cnt
1	0.37	-	-	-	0.13	0.20	-	0.29	0.01	-
2	0.45	0.07	-	-	-	0.17	-	0.30	-	0.01
3	0.39	-	-	-	-	0.20	-	0.32	0.04	0.05
4	0.30	-	-	-	0.03	0.20	0.05	0.33	0.03	0.07
5	0.42	-	-	0.05	0.17	0.05	-	0.32	-	-
6	0.36	-	-	0.06	0.12	0.11	-	0.31	0.02	0.01
7	**0.31**	**-**	**-**	**-**	**0.21**	**0.17**	**-**	**0.31**	**-**	**-**
8	0.46	-	-	0.02	-	0.19	-	0.30	0.03	-
9	0.20	-	-	0.01	0.20	0.20	-	0.16	0.15	0.07
10	0.26	0.07	-	0.08	0.21	0.13	-	0.19	0.07	-
11	0.26	-	-	0.04	0.24	-	-	0.26	0.08	0.12
12	0.16	-	-	0.04	0.21	0.16	0.05	0.21	0.06	0.11
13	0.25	-	-	0.04	0.24	0.06	-	0.25	0.04	0.12
14	0.16	0.02	-	0.04	0.21	0.16	-	0.14	0.12	0.16
15	0.20	-	-	0.06	0.20	0.18	-	0.20	0.06	0.10
16	0.16	0.01	-	0.01	0.22	0.16	0.03	0.19	0.12	0.10
17	0.18	-	-	0.10	0.18	0.16	-	0.18	-	0.18
18	0.16	-	-	0.04	0.22	0.17	0.05	0.14	0.12	0.11
19	0.22	0.01	-	0.04	0.22	0.16	0.05	0.09	0.12	0.10
20	0.21	-	-	0.03	0.21	0.15	0.04	0.16	0.11	0.10
21	0.26	0.02	-	0.06	0.07	0.20	0.06	0.18	0.15	-
22	0.26	-	-	0.03	0.16	0.14	-	0.26	0.03	0.12
23	0.19	0.12	-	0.03	0.17	0.17	0.05	0.15	0.12	-
24	0.18	0.14	-	0.03	0.18	0.13	-	0.12	0.10	0.13

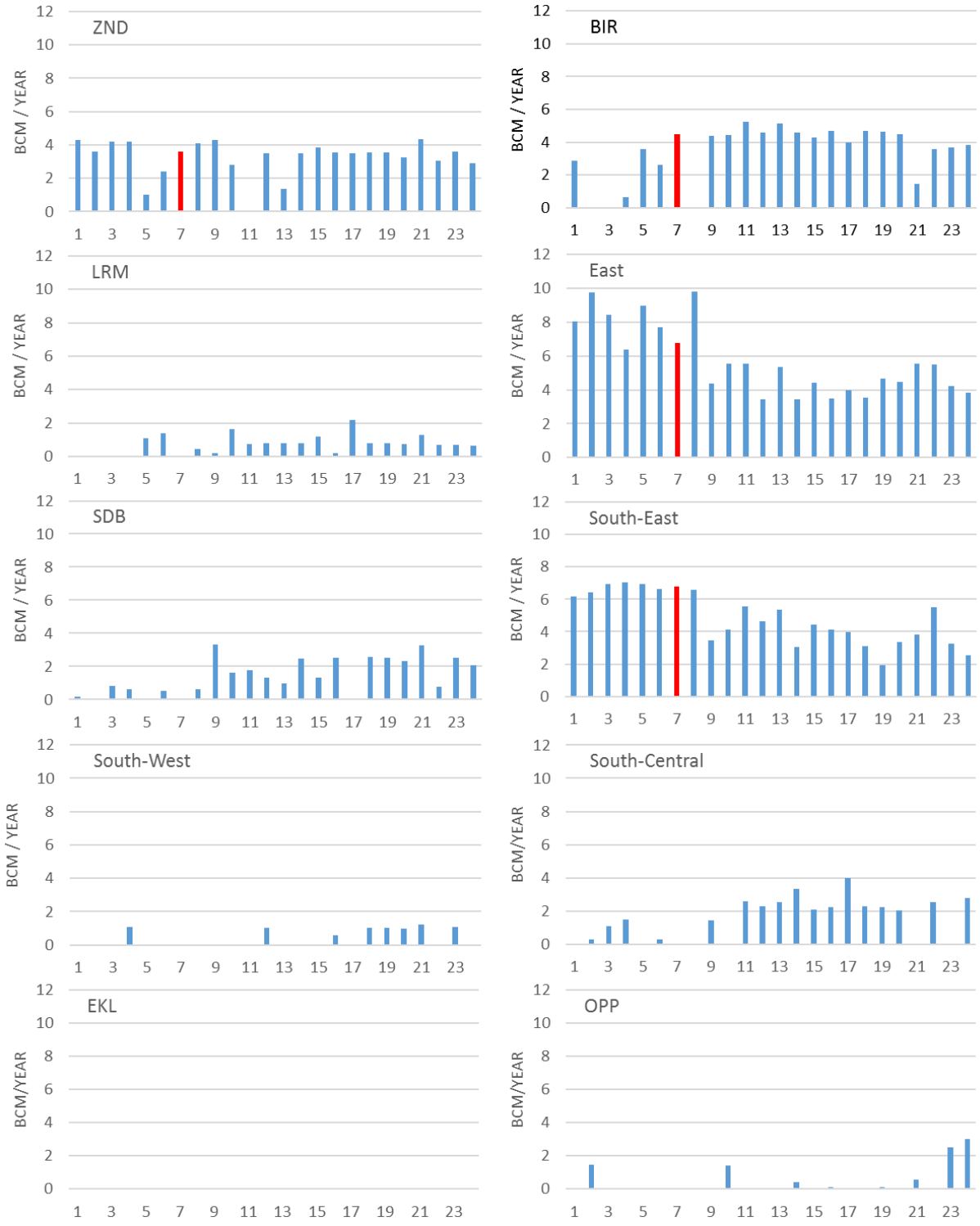


Figure D-7: Optimised production splits for different logic tree branches for population weighted PGV metric. Branch 7, indicated in red, gave the optimal solution when evaluated in terms of the logic tree mean.

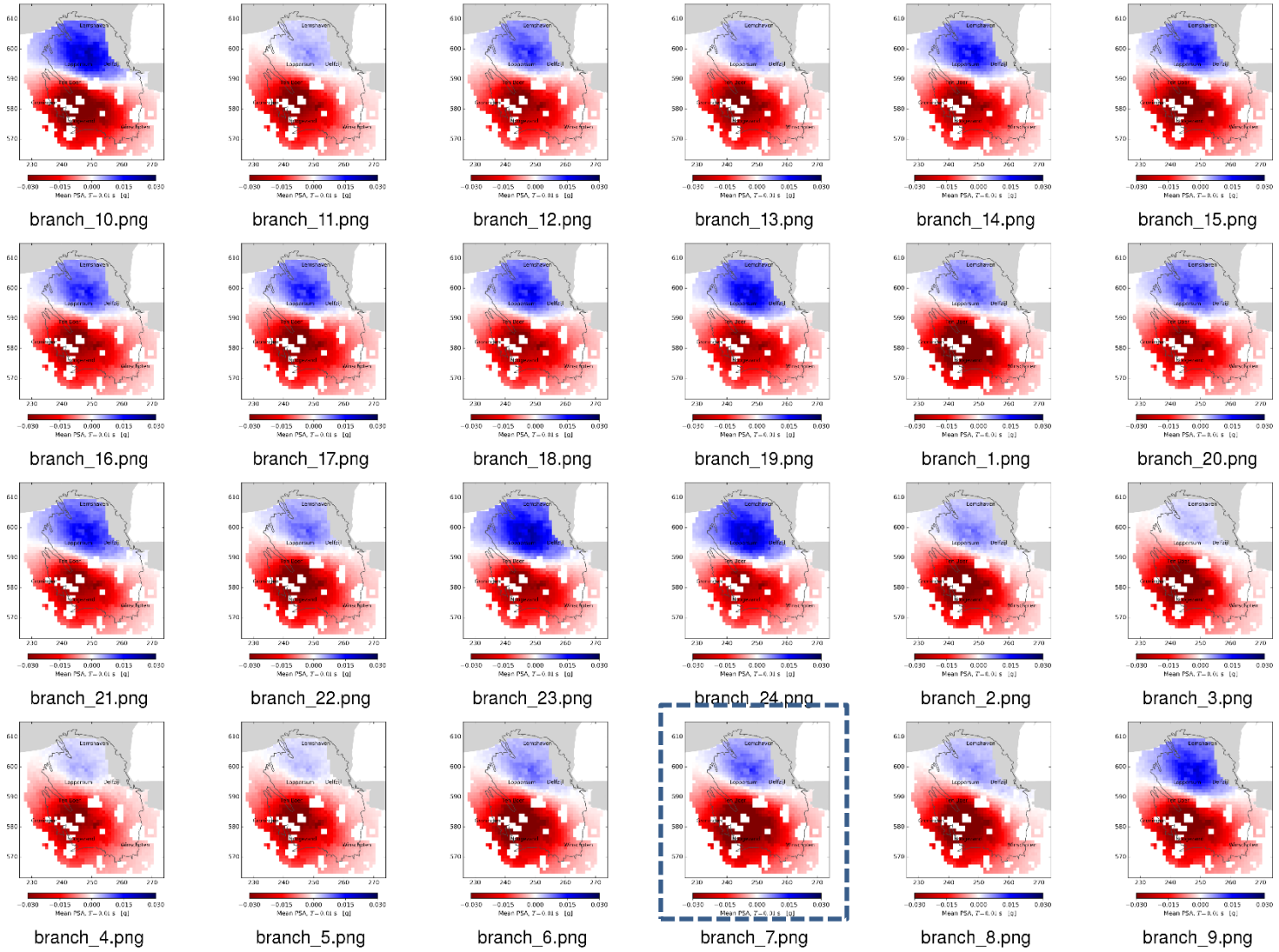


Figure D-8 Mean PGA hazard difference maps (scaled [-0.03,0.03]) for Population-weighted PGV driven optimisation, at average 0.2% annual chance of exceedance (1 in 475 years) from 1/1/2018 to 31/12/2022 (optimal solution indicated)

D.5 Hybrid metric

Table D-9: Seismic metrics for optimised distribution based on hybrid objective. Percentage improvement is w.r.t. BP17 reference case.

Branch	pwPGV	maxPGA	maxPGV	# EQ	pwPGV	maxPGA	maxPGV	# EQ
1	**0.043**	**0.228**	**0.138**	**95**	**-16%**	**-10%**	**-4%**	**-17%**
2	0.044	0.227	0.139	96	-15%	-10%	-3%	-16%
3	0.044	0.227	0.139	97	-15%	-10%	-3%	-16%
4	0.044	0.226	0.137	96	-15%	-10%	-4%	-17%
5	0.044	0.227	0.136	96	-14%	-10%	-5%	-17%
6	0.044	0.226	0.138	96	-15%	-10%	-4%	-17%
7	0.044	0.227	0.139	95	-15%	-10%	-3%	-17%
8	0.044	0.227	0.137	96	-15%	-10%	-4%	-17%
9	0.044	0.230	0.140	97	-14%	-9%	-2%	-16%
10	0.044	0.228	0.138	96	-14%	-9%	-3%	-16%
11	0.045	0.228	0.137	97	-13%	-10%	-4%	-15%
12	0.044	0.226	0.138	96	-15%	-10%	-4%	-16%
13	0.045	0.229	0.139	98	-13%	-9%	-3%	-15%
14	0.044	0.227	0.138	96	-15%	-10%	-3%	-17%
15	0.044	0.228	0.139	96	-15%	-10%	-3%	-17%
16	0.044	0.228	0.138	96	-15%	-10%	-3%	-17%
17	0.044	0.230	0.139	98	-14%	-9%	-3%	-15%
18	0.044	0.229	0.138	97	-15%	-9%	-4%	-16%
19	0.044	0.228	0.138	97	-14%	-10%	-4%	-16%
20	0.044	0.228	0.137	95	-15%	-10%	-4%	-17%
21	0.044	0.231	0.141	98	-14%	-8%	-2%	-15%
22	0.044	0.229	0.138	96	-15%	-9%	-3%	-16%
23	0.044	0.228	0.139	96	-15%	-10%	-3%	-16%
24	0.044	0.229	0.138	97	-14%	-9%	-3%	-16%

Table D-10: Optimised normalized production splits for different logic tree branches for hybrid objective.

Branch	East	OPP	EKL	LRM	BIR	ZND	S-West	S-East	SDB	S-Cnt
1	**0.48**	**-**	**-**	**-**	**0.19**	**-**	**-**	**0.32**	**-**	**-**
2	0.36	-	-	-	0.22	0.02	-	0.33	0.01	0.07
3	0.32	-	-	-	0.22	0.06	-	0.32	-	0.07
4	0.36	-	-	-	0.21	0.01	-	0.32	0.01	0.09
5	0.35	-	-	-	0.20	-	-	0.32	-	0.14
6	0.37	-	-	0.01	0.22	-	-	0.34	0.01	0.05
7	0.48	-	-	-	0.17	-	-	0.33	-	0.03
8	0.40	-	-	-	0.20	-	-	0.32	0.01	0.07
9	0.30	-	-	-	0.23	0.06	-	0.33	0.02	0.06
10	0.32	-	-	-	0.23	0.02	-	0.32	0.01	0.09
11	0.28	-	-	-	0.23	-	0.04	0.32	0.01	0.11
12	0.33	-	-	-	0.23	0.04	-	0.33	-	0.07
13	0.31	-	-	-	0.24	0.01	0.04	0.31	-	0.09
14	0.36	-	-	-	0.23	-	-	0.32	-	0.08
15	0.35	-	-	-	0.21	-	-	0.33	0.06	0.05
16	0.35	-	-	-	0.23	0.01	-	0.33	0.01	0.08
17	0.47	-	-	-	0.11	-	-	0.30	0.11	-
18	0.45	-	-	-	0.09	0.02	-	0.32	0.08	0.04
19	0.45	-	-	0.02	0.07	0.05	-	0.32	0.05	0.04
20	0.49	-	-	-	0.16	-	-	0.31	-	0.04
21	0.46	0.02	-	-	0.13	-	-	0.32	0.08	-
22	0.51	-	-	0.01	0.10	-	-	0.32	0.07	-
23	0.48	-	-	-	0.09	0.06	-	0.33	0.04	-
24	0.49	-	-	-	0.09	-	-	0.31	0.09	0.02

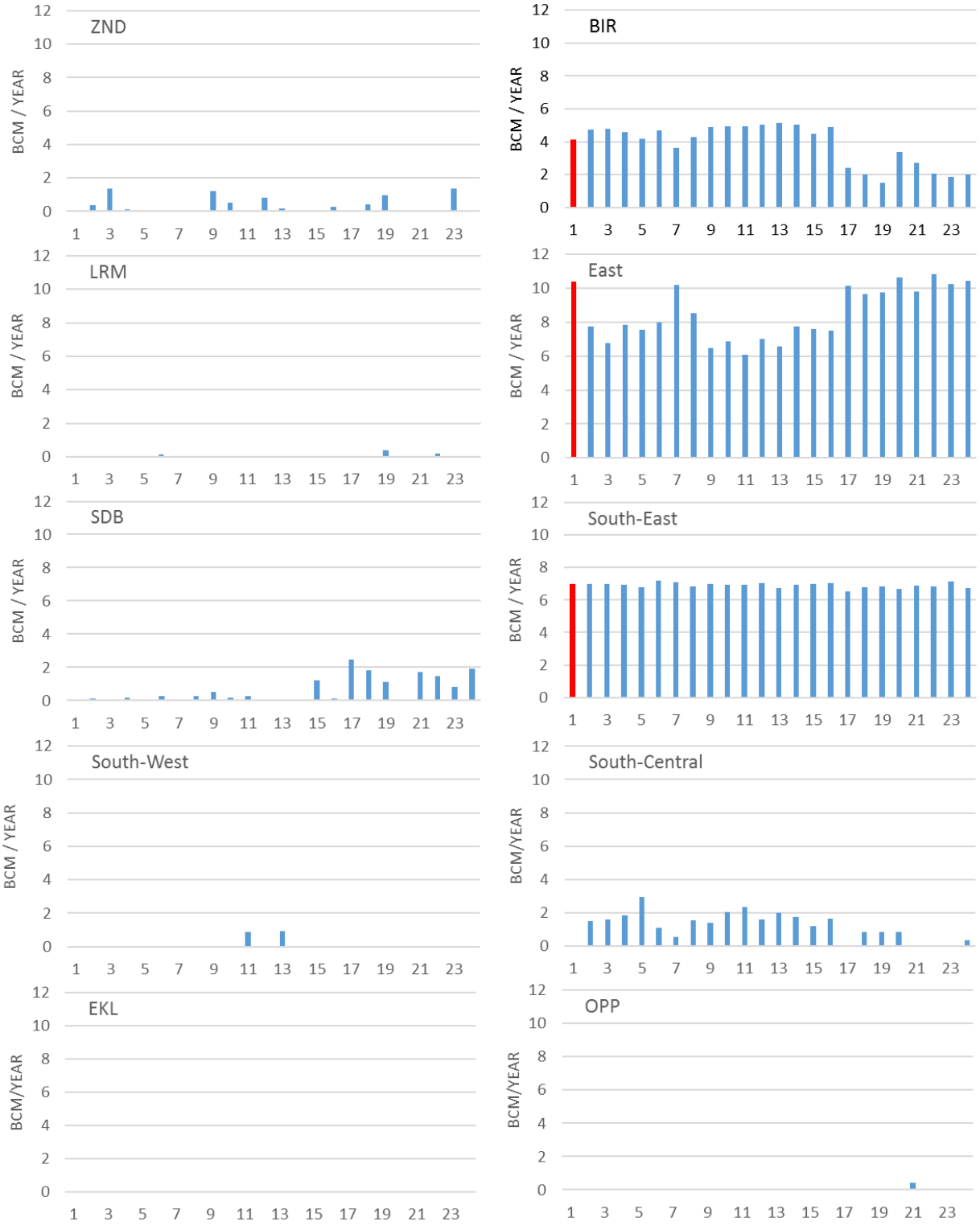


Figure D-9: Optimised production splits for different logic tree branches for hybrid metric. Branch 1, indicated in red, gave the optimal solution when evaluated in terms of the logic tree mean.

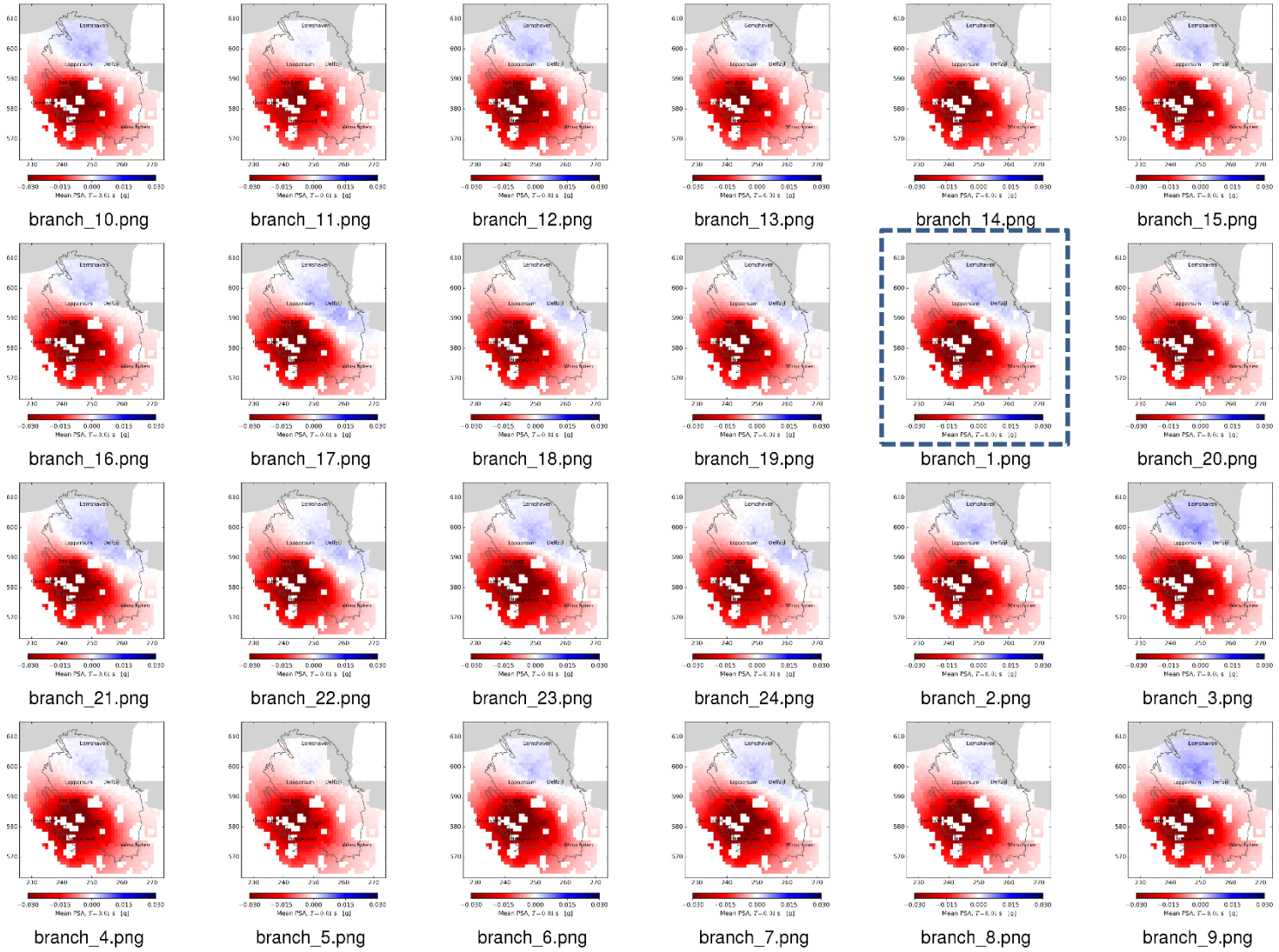


Figure D-10 Mean PGA hazard difference maps (scaled [-0.03,0.03]) for Hybrid (EventCount+PWPGV) driven optimisation, at average 0.2% annual chance of exceedance (1 in 475 years) from 1/1/2018 to 31/12/2022 (optimal solution indicated)

Appendix E Operational limitations

The Groningen production system was largely designed and constructed in the 1960's. The design was focussed on the highest capacity and reliability during periods of high demand (winter), and did not anticipate any operational requirements with respect to management of production induced seismicity. In this chapter the components making up the Groningen system are described, along with recent changes in operating philosophy resulting from the requirement to manage induced seismicity, providing background for potential further adjustments to reduce seismic hazard and/or risk.

E.1 Link to the gas market

The Groningen production system is part of an integrated system in Northwest Europe that supplies gas into a market of highly variable demand (Figure E-1). Gas is transferred from NAM's Groningen system to the system of GasUnie Transport Services (GTS) across seven "Overslagen" (custody transfer stations, or OV's). Offtake distribution per OV is controlled by GTS by manipulation of the pipeline pressure and taking into account the actual market demand. Given that the entire Groningen quality gas market has evolved around the Groningen field, the field forms the starting point for the GTS infrastructure and sits at a cross-roads of the GTS pipeline network. Consequently, any change in operating the Groningen ring potentially impacts GTS operations and its ability to re-distribute gas within its own existing network. Traditionally GTS relied on Groningen ring to supply gas into its various pipeline networks towards the different gas markets. On a high level, there are three main gas markets:

- Towards West Netherlands
- Towards Germany
- Towards South Netherlands/Germany/Belgium/France

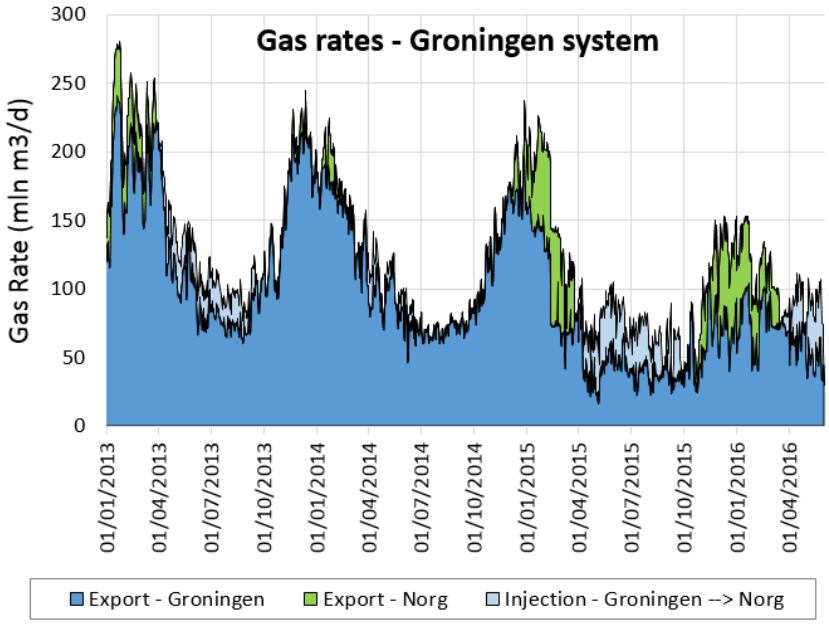


Figure E-1 Production fluctuations over the period from 2013

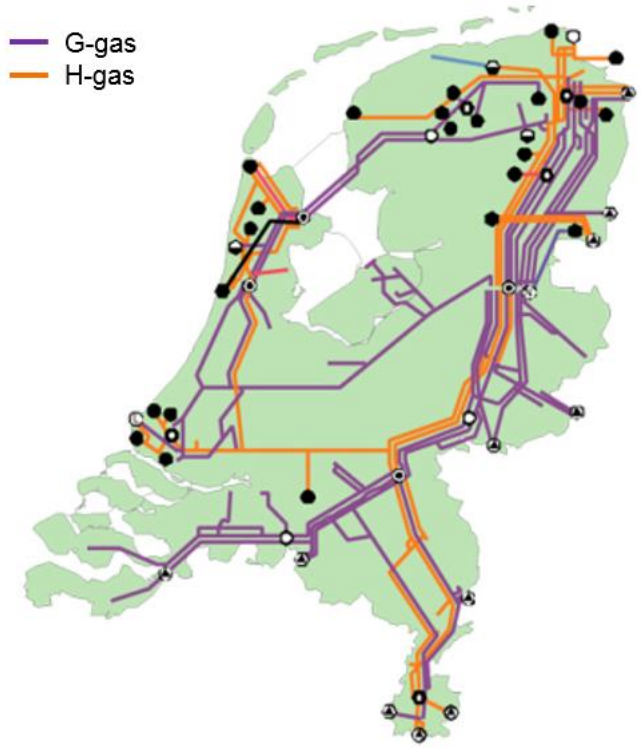


Figure E-2 Gasunie Transport Solutions pipeline grid in the Netherlands. All L-gas pipelines originate from the Groningen field.

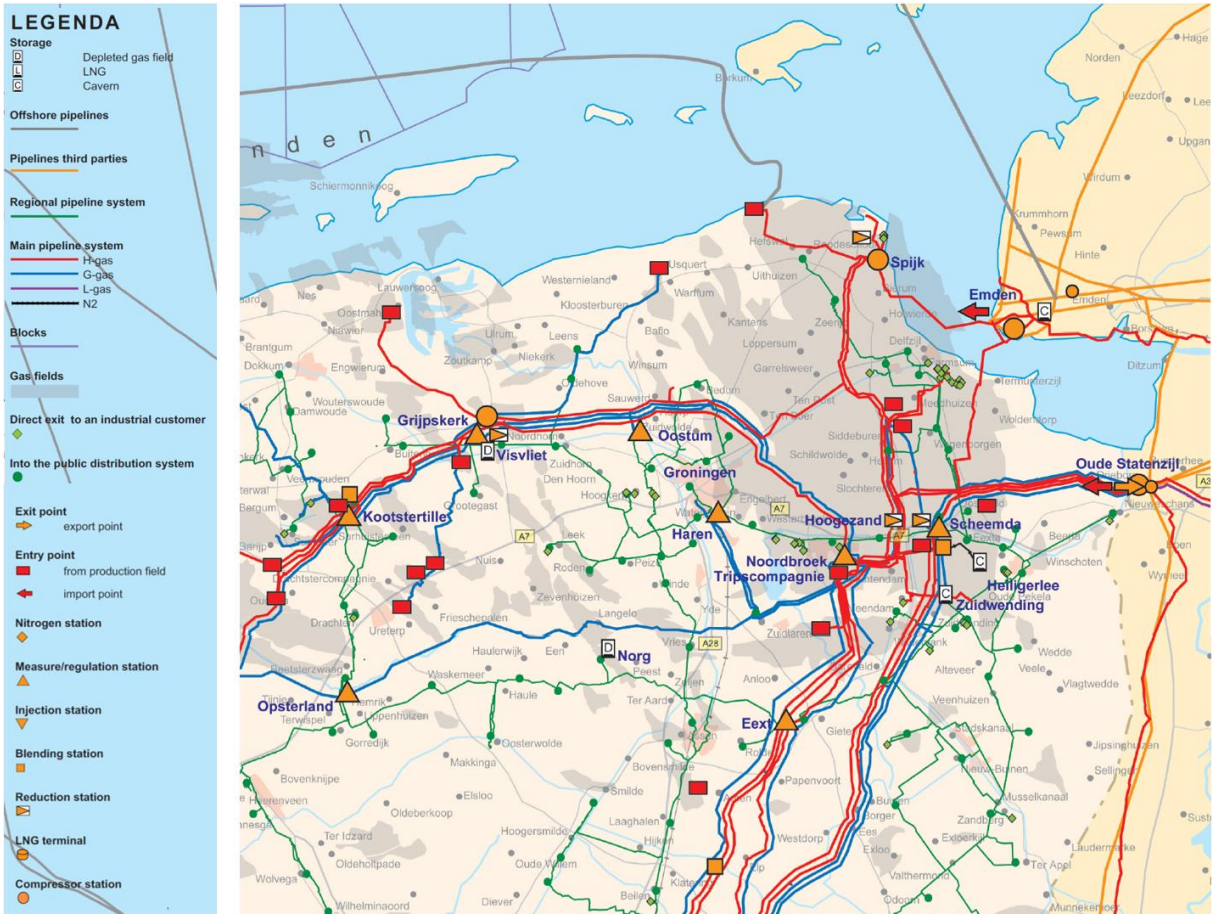


Figure E-3 A closer look at the Gasunie Transport Solutions pipeline grid around Groningen (source: GTS).

Based on the Grid Connection Agreement between NAM and GTS, gas leaving the Groningen ring at the OV's into the GTS grid needs to meet the following specifications[12]:

- Pressure specification between 55 and 65 bar
- Wobbe Index between 43.46 to 44.41 MJ/Nm^3.
- Contaminant Limitations.

The custody transfer stations at Oude Statenzijl (OSZ) serves the German market and is operated by GTS. As GTS is not able to manipulate the flow from OSZ, it relies on NAM to supply at the higher end of the contractual pressure window to ensure delivery to the North-German gas market downstream of OSZ.

E.2 Ring System and Underground Gas Storage

The Groningen production clusters and the custody transfer stations (OV's) are interconnected by a pipeline network that roughly makes up a ring. This setup allows for a very high operational flexibility, as any cluster can ultimately flow over any OV. Due to the mutual distances some configurations are more practical than others, as there may be up to 10 bar pressure drop involved.

[12] These specifications are described in Article 11 and 12 of the Gaswet

Due to the 'Instemmingsbesluit' requiring to keep Groningen production flat, there is an increased utilization of/dependence on the Norg Underground Gas Storage (Figure E-1). Norg is produced in winter, and re-filled with Groningen gas in summer. Thus Norg acts as a capacity provider, which can be used to keep the demand on the Groningen field much more constant throughout the seasons.

It was foreseen the UGS Norg would be filled via the dedicated NorGron pipeline (OV-SAP to Norg) using the central located production clusters. However with the restriction on most of these clusters, taking into account the most flexible utilisation of the Groningen ring, the UGS is mainly filled with Southern clusters, making use of splitting the ring in a high pressure and low pressure section, and a number of flow control valves. Alternative options to feed UGS Norg are feasible but would require shut-in of one or more OV's and consequently lead to an impact in GTS operations.

In case GTS would be able to accommodate custody transfer over only a limited number of OV's, relatively simple set-ups of the ring split would be possible which provide a high operational flexibility, see Figure E-4. To date, more complicated splits are used which provide less operational flexibility, see Figure E-5 and Figure E-6.

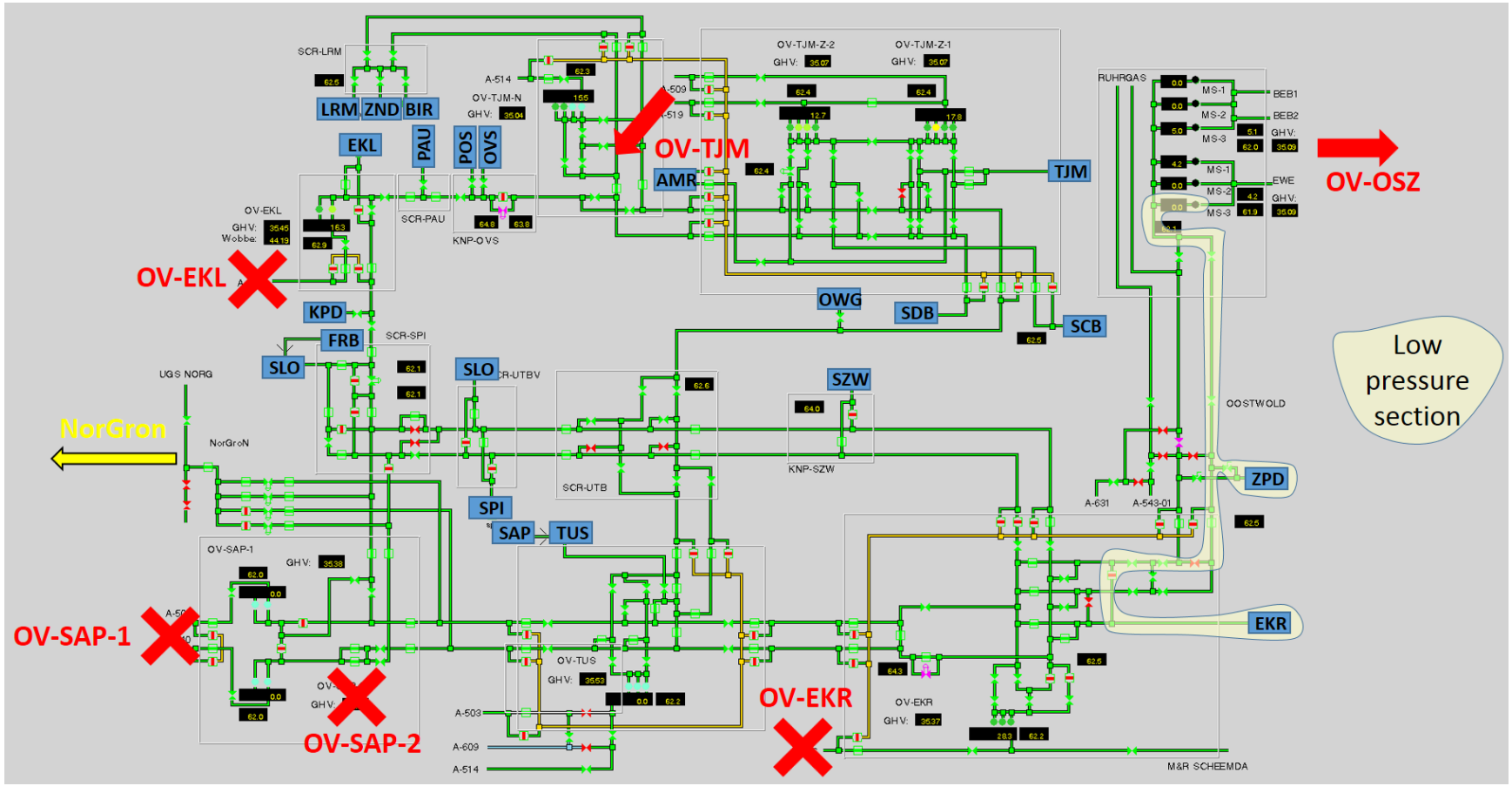


Figure E-4: HP/LP ring split configuration for a relatively simple ring-split. The low-pressure section is highlighted in a yellow cloud, whereas the remainder of the ring would be at a higher pressure.

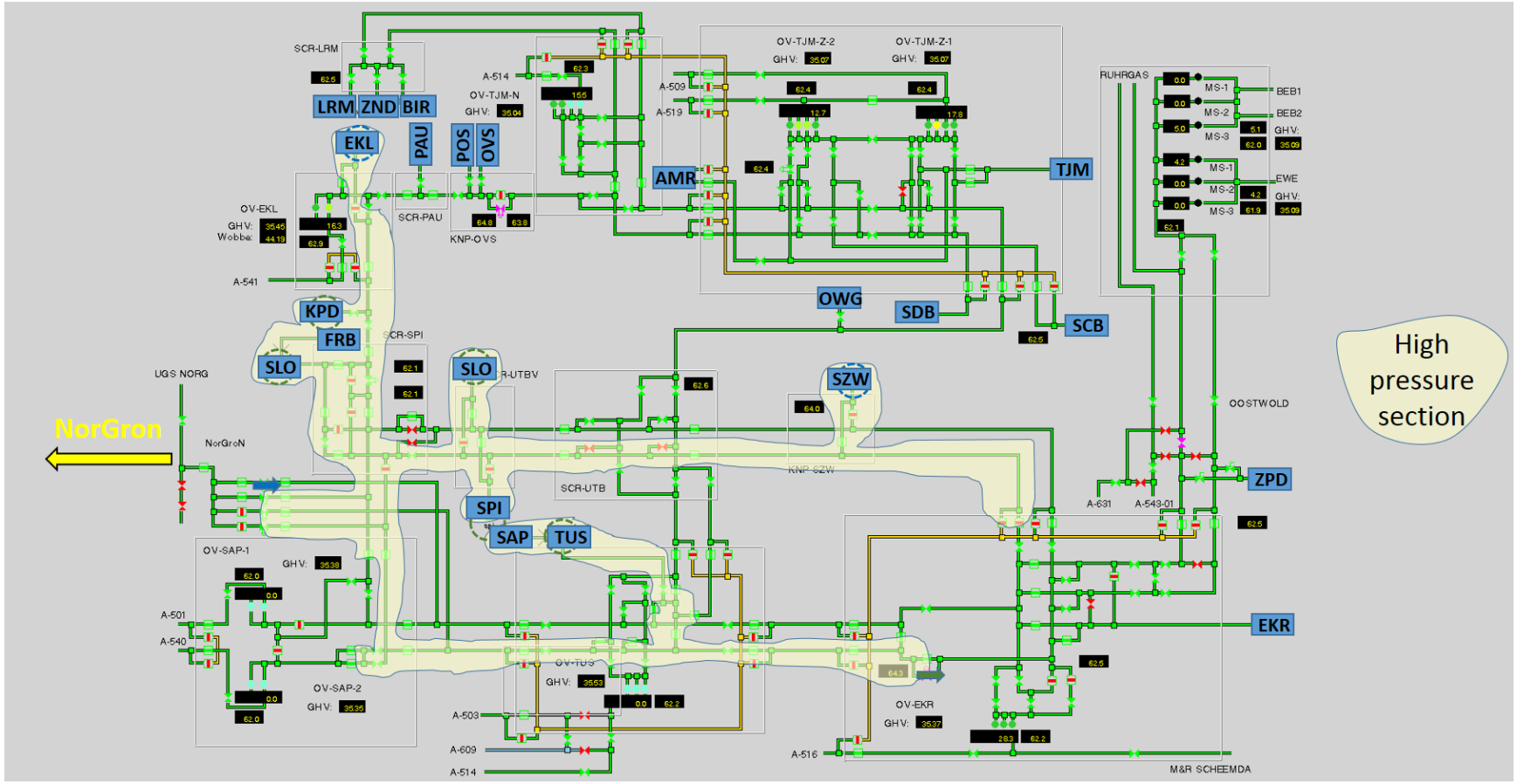


Figure E-5: First example of a currently operationally feasible pressure split in the ring. In this case, the yellow cloud indicates the section of the ring that would be operated at an increased pressure, whereas the remainder of the ring would be at a lower/normal pressure.

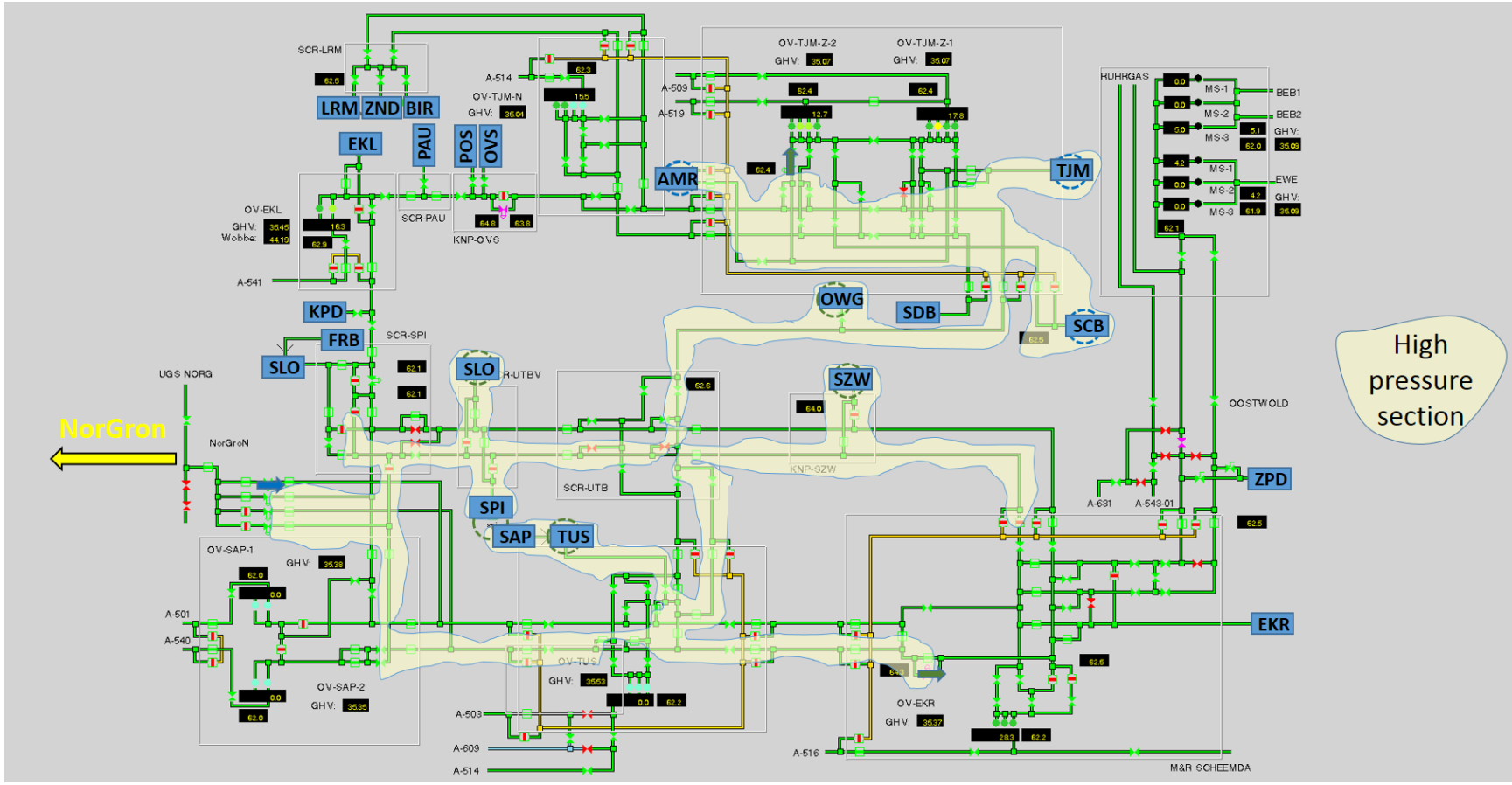


Figure E-6: Second example of a currently operationally feasible pressure split in the ring. Again, the yellow cloud indicates the section of the ring that would be operated at an increased pressure, whereas the remainder of the ring would be at a lower/normal pressure.

E.3 Gas Production Wells

Due to the excellent reservoir quality, typically the Groningen wells are highly productive. However, with depleting reservoir pressure the well capacities are steadily declining.

E.4 Clusters and Compressors

The current Groningen production system comprises 20 production clusters, which are all on compression. All clusters are on first stage compression (B-bundle cartridge[13]), with the exception of:

- Schaapbulten (SCB), 2nd stage
- Eemskanaal (EKL), 1st stage with an A-bundle cartridge

The 22 clusters include two tie-backs (no dedicated compressor):

- Froombosch (FRB) is a tie-back to Slochteren (SLO)
- Sappemeer (SAP) is a tie-back to Tusschenklappen (TUS)

A schematic representation of the main components making up a production cluster are given in Figure E-7. Figure E-8 gives an overview of the actual set-up in the field. As can be seen in Figure E-7, in a steady operation all components run within a certain pressure and temperature domain. A so-called cold start-up implies a start-up period of several hours up to several days before the full cluster process train has reached stable operations within the required operating envelopes. This start-up period will depend on the duration of the preceding production stop, the ambient temperature, and the total number of locations involved in the start-up to produce at a certain production flow.

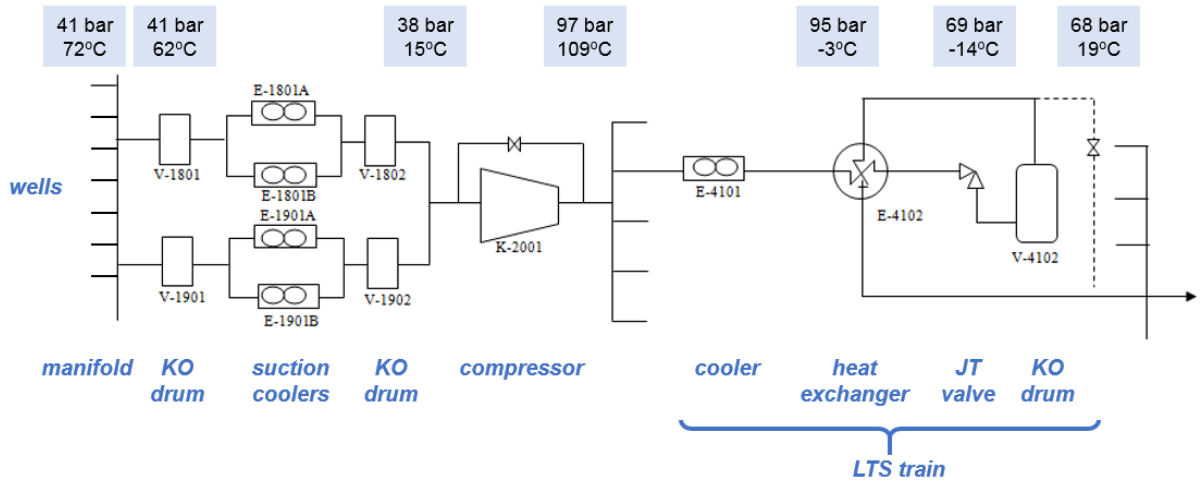


Figure E-7 Simplified process diagram for a Groningen production cluster

[13] The difference between the A- and B-bundle is the configuration of the rotor (8 blades versus 5 blades).

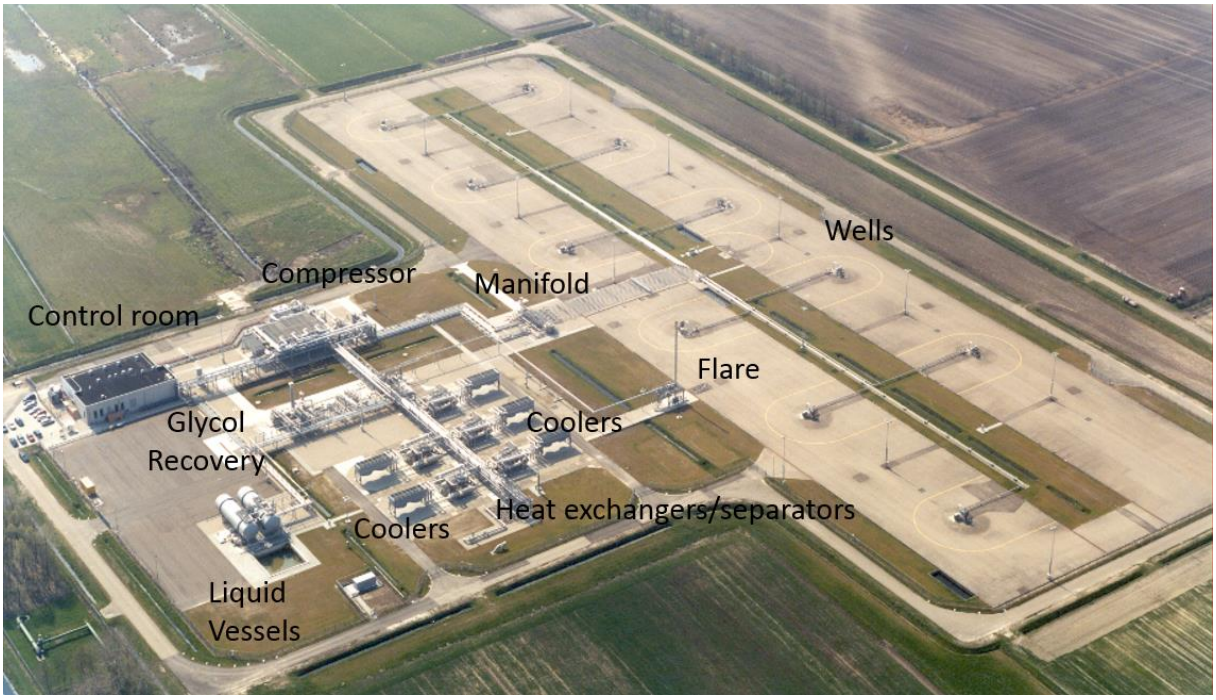


Figure E-8 Overview of a production cluster location.

E.5 Conclusion operational constraints

The Groningen production system is currently operated outside the envelope anticipated during its design and construction. The interplay of the many components and requirements make the operation of the Groningen system therefore a complex task. Over the last years' experience has been gained in operating the Groningen gas production system outside the originally intended operating envelope. Although not all combinations of requested offtake at the overslagen have been tested, with this experience a more flexible operating philosophy of the network has been implemented. In order to investigate which additional adjustments in the surface system may lead to a further reduction in the seismic hazard or seismic risk, NAM and GTS will need to align on requirements for Groningen delivery across OV's. Possibly field tests are required for GTS and NAM to test the operating envelopes.

Appendix F Post-processing using Machine Learning

F.1 What is meant by Machine Learning

In Reference [24] a popular definition of a Machine Learning (ML) algorithm is given. It is said to:

> *"learn from experience e with respect to some class of tasks t and performance measure m if its performance at tasks in t, as measured by m, improves with experience e".*

In the context of this report, this would translate as:

experience e	control response combinations from the experimental design,
task t	predicting the output for the objective of choice (population weighted PGV, max PGV, max PGA, number of tremors, or hybrid thereof)
m	error metric, such as the Mean Square prediction Error (MSE).

F.2 Measuring the Predictive Performance of a Model

Mathematical Notation and Definitions:

The mathematical notations and abbreviations used are as follows:

n	number of controls, i.e. the number of cluster groups that can contribute to the total production ($n > 0$)
p	number of available control response data points ($p > 0$)
c_{ij}	fractional contribution of cluster i for branch j to the normalized production of 1, with $1 \leq i \leq p$ and $1 \leq j \leq n$; $c_{ij} \in [0,1]$

Furthermore, we denote a response of the Mores-HRA coupling (e.g. pwPGV) by:

$$r_i \in \mathbb{R}$$

if the corresponding controls $c_i \in [0,1]^n$ are used as input. Note that for the sake of simplicity the actual units are omitted since they should be clear from the context.

To abbreviate our notation, we will refer to the i-th data point, consisting of control settings and a Mores-HRA response, as:

$$x_i = (c_i, r_i) \in [0,1]^n \times \mathbb{R}, \qquad \text{with } 1 \leq i \leq p.$$

We introduce the following notations :

$$R = (r_i, \dots, r_p), \text{ and } X = (x_1, \dots, x_p).$$

Let $T \subseteq X$ be a subset of data points in X. We denote by $|T|$ the number of points in the set. Furthermore, we denote a proxy model that has been trained/calibrated on the points in X that maps the controls to a response of the Mores-HRA coupling by

$$f_T: [0,1]^n \mapsto \mathbb{R}$$

For a control point $c_i \in [0,1]^n$ we denote the response of the proxy model, that was trained on T, by:

$$\tilde{r}_i = f_T(c_i).$$

In analogy to the previous abbreviations, we let

$$\tilde{R} = (\tilde{r}_i, \dots, \tilde{r}_p).$$

Finally, we call a function that allows to compare true Mores-HRA responses with predicted Mores-HRA responses (through a proxy model) as an error metric:

$$m: \mathbb{R}^p \times \mathbb{R}^p \mapsto \mathbb{R}$$

General Remarks:

In order to calibrate the proxy model, it is "trained" on a subset of X, which we will call the training set:

$$T \subseteq X.$$

The performance of the model is measured on a separate subset of X, the so-called validation set:

$$V \subseteq X.$$

When measuring the predictive performance of any model in an error metric m, it is important to have a methodology in place that can estimate the performance of the model on as-of-yet unseen data.

If the performance of a model is assessed *in sample*, there is overlap between the training and test data sets, such that:

$$T \cap V \neq \emptyset,$$

For an in-sample assessment, the error metrics are to a certain extent only statements about how well the existing training data can be reproduced.

For models that provide sufficient degrees of freedom in their internal representations, there is always the danger of over-fitting to T. In general, with increasing model flexibility and complexity the prediction error measured in sample tends to get smaller. However, this does not imply that the generalization error of the model on data that was not used to train the model (such as points in $X \setminus T$) will decrease as well. Hence, one commonly accepted way of measuring and reporting the predictive power of a model is to test it only on unseen data. This is termed *out of sample validation*.

Practically speaking, for out of sample validation the data set X is partitioned into training and validation data sets T and V, normally in a ratio of around 80:20 or 90:10. The model is then trained on the respective training data set T and used to predict the data in the corresponding validation data set V. The performance metrics are then computed on the known responses of the validation data (from the Mores-HRA coupling) against the prediction by the model f_T. Additionally, uncertainties in the error measure itself are quantified and reported in terms of the standard error, which is the associated standard deviation.

It is good practice to not only randomly partition the data set once in training and validation set, but to repeat this procedure several times. A common scheme for subdividing the data into equally sized partitions is *cross-validation*, which is described in more detail in the following.

Estimating the Generalization Error and Associated Standard Error via Cross Validation:

The general objective, as outlined above, is to compute the generalization error and associated confidence interval of a model f_T on unseen data V in the chosen error metric(s). It is important to realize that the choice of the partition in which the given data is split into training and test set will have an impact on both the trained model f_T and on the error metrics computed on the test set. In order to reduce this effect related to the partitioning of the input data, a common strategy is to run several model training and prediction experiments on different partitions of the data and to report the mean of the error metrics and the associated standard errors. If the experiment is performed l times, the uncertainty in the estimates is reduced by a factor of $\frac{1}{\sqrt{l}}$.

A common partitioning technique is to split the given data into l (almost) equally sized bins which contain $\left\lceil \frac{p}{l} \right\rceil$ elements, except potentially the last bin which may have less elements. We abbreviate the data points in the i-th bin with B_i. Then consequently l training and test sets are formed where the i-th test set is given by B_i and the i-th training set is given by $X \setminus B_i$. In total l models are trained and evaluated. In the end, the mean of the individual error metrics is reported and the associated standard error is computed. This procedure is known as *cross-validation*. A nice property of this scheme is that it ensures that every point in X is predicted exactly once out of sample. A schematic overview of 5-fold cross-validation is contained in Figure F-1.

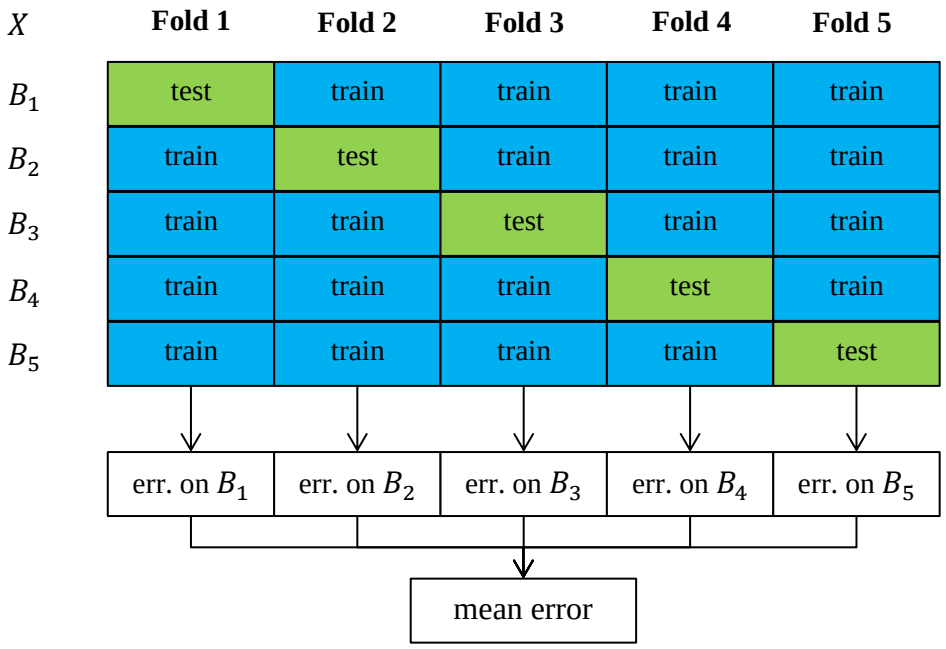


Figure F-1: Stylized example of 5-fold cross-validation

Error Metrics

The following error metrics are used related to the estimation of the predictive performance of the proxy models:

- The Mean Squared Error of a model f_T, with respect to training set T and validation set V, is defined as:

$$\mathrm{MSE}(f_T, V) = \frac{1}{|V|}\sum_{i=1}^{|V|}(r_i - \tilde{r}_i)^2,$$

- The Root Mean Squared error is defined as:

$$\mathrm{RMSE}(f_T, V) = \sqrt{\frac{1}{|V|}\sum_{i=1}^{|V|}(r_i - \tilde{r}_i)^2}\,.$$

Note that in these error measures larger deviations in predictions from the true values get penalized more than smaller deviations.

The associated Standard Errors are given by:

- For MSE:

$$\mathrm{SE}_{\mathrm{MSE}(f_T,V)} = \frac{\mathrm{sd}\left(\left(R - \tilde{R}\right)^2\right)}{\sqrt{|V|}}$$

- For RMSE:

$$\mathrm{SE}_{\mathrm{RMSE}(f_T,V)} \approx \frac{\mathrm{SE}_{\mathrm{MSE}(f_T,V)}}{2 \cdot \mathrm{RMSE}(f_T, V)} = \frac{1}{2}\frac{\mathrm{sd}\left(\left(R - \tilde{R}\right)^2\right)}{\sqrt{\sum_{i=1}^{|V|}(r_i - \tilde{r}_i)^2}}$$

where $\mathrm{sd}(\cdot)$ denotes the usual standard deviation and $\left(R - \tilde{R}\right)^2 = \left((r_1 - \tilde{r}_1)^2, \ldots, \left(r_{|V|} - \tilde{r}_{|V|}\right)^2\right)$.

F.3 CART Trees and Random Forests

A Classification and Regression Tree (CART) is a tree-based ML method. Tree based methods partition the feature space into a set of (high-dimensional) rectangles and fit a simple model for each rectangle, Reference [25]. We mention this method herein because it is important for understanding the Random Forest algorithm that is the primary ML algorithm we use for the analysis of the Mores-HRA coupling.

To construct a CART decision tree, first a root node is created. From the list of features, which in our situation are the cluster offtake splits, an optimal feature for branching and an associated splitting value

are determined based on the minimization of the sum of the square errors between the prediction of the resulting model and the actual value of the prediction target. This branching feature and value is used to partition the training data into a "left" and "right" subset. The left and right branches correspond to samples that have a value less than or equal to the splitting value and greater than the splitting value, respectively, for the feature upon which the splitting is applied. For each of the branches, a new child node is defined and the previous procedure is repeated, based on the subset of the data that remains in the branches in question. It is generally considered to be sensible to use a binary split, at each branch, primarily since it is easier to implement but also because in some situations not enough data would remain in each of the branches, in the next level down in the tree. The splitting procedure is applied, recursively, until only one data point remains in each of the sub-branches. A terminal node is then defined, such that the value of the prediction target is stored. This is the value that the tree would predict for a new sample that follows the same route through the tree. In practice, for a single CART tree, it would be common to apply pruning, a method in which the terminal nodes contain more than one data point, for which an average would be taken. This avoids the tendency to over-fit to the training data by trading-off bias and variance. Note that the CART trees which are used inside the Random Forest algorithm are not pruned per-se, but they are also not grown to the point at which each termination node is a single data point. An ensemble of trees is used to reduce the chance of over-fitting. An illustrative example of an individual CART tree can be seen in Figure F-2. Internal nodes are represented by white circles and terminal nodes by red circles. Suppose that you have 10 controls and $c_3 = 0.1$, then the prediction for pwPGV from this individual tree is 0.06.

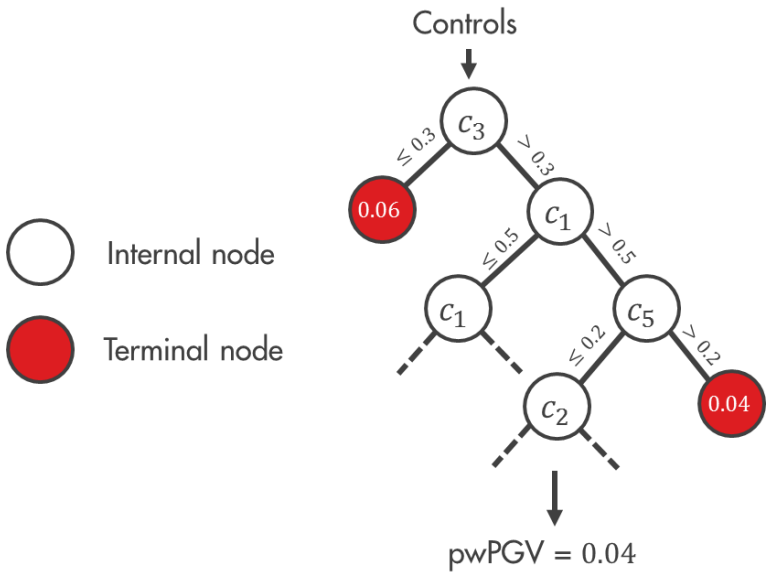


Figure F-2: Illustrative example of a CART tree that relates controls to a Mores-HRA response

The Random Forest algorithm uses an ensemble of CART trees. It was originally proposed by Breiman in 2001 and has since been cited more than 30000 times [26]. This illustrates its widespread use and acceptance throughout academia and in many industrial applications. Each of the CART trees is constructed on a random sub-sample of the training data and using only a subset of the features, chosen, again, at random. The prediction of the model is obtained by traversing each individual CART tree with

the test data and in the context of regression, taking the mean of the predictions of each individual CART tree. An illustration of a Random Forest model, consisting of several hundred individual CART trees, can be seen in Figure F-3. The main tuning parameters are the number of trees and the number of features that are randomly sampled. The Random Forest is typically seen to be insensitive to hyper-parameter tuning and works well with default settings. It is good at capturing non-linear relationships in small to medium sized data sets but is considered too computationally expensive for "Big Data". It can handle both categorical and numerical data and is essentially invariant to monotonic transformations of the features. All predictions it makes remain within the range of the observations in the training data. We use the implementation available in the '*randomForest'* package for R, Reference [27]. Note that the programming language R is the de facto standard for statistical computing and contains many implementations that have been checked numerous times for statistical rigor.

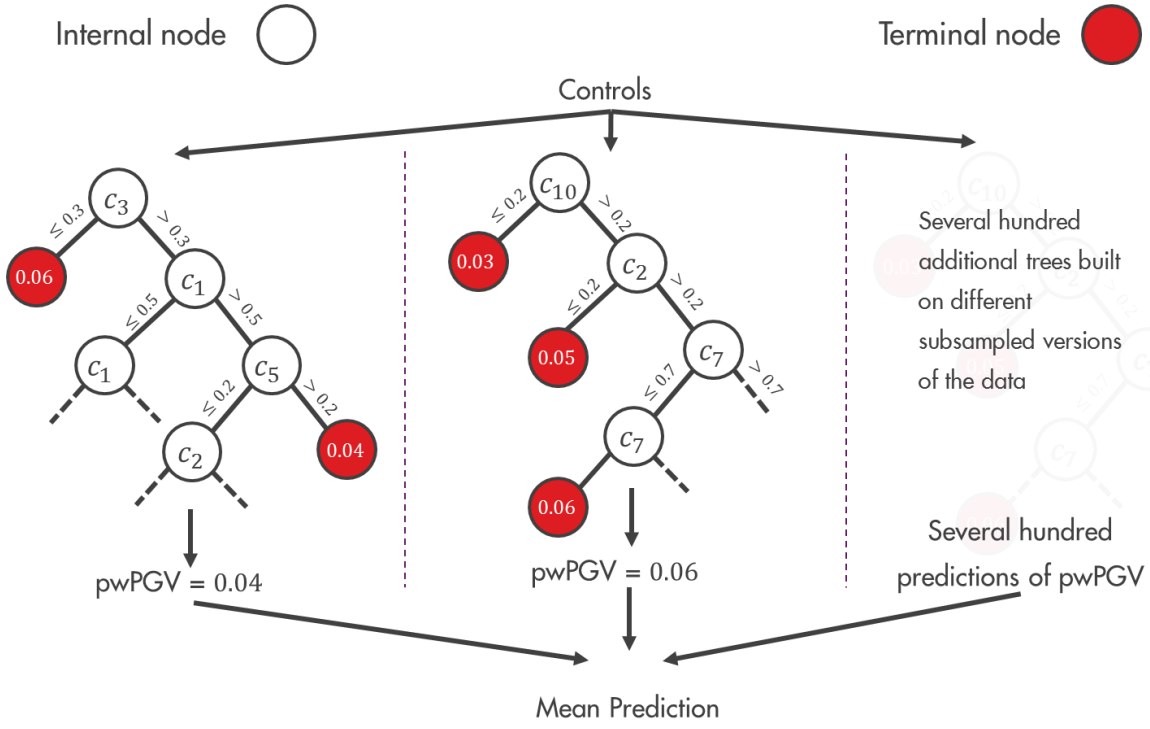


Figure F-3: Illustrative example of a random forest model, consisting of an ensemble of CART trees.

F.4 Relative Variable Importance Analysis

It is not our sole intention to build a proxy model for the relationship between the cluster offtake splits and the prediction target of choice, e.g. population weighted PGV. We note however, that by using the Random Forest algorithm, it is possible to construct a proxy model with an out of sample coefficient of determination of over 0.95. This would generally be considered sufficiently accurate such that a variable importance assessment, performed on the proxy model, would be informative.

Our primary objective is to use the proxy model to understand the relative impact of changing each of the cluster offtake splits on the prediction target in question. In the simplest case, one could consider

constructing a Pareto plot by systematically varying each of the cluster offtake splits in a univariate manner, adjusting each of the other offtake rates, proportionally. However, this analysis does not account for the interactions and subsequent impact on reservoir response of a multivariate setup. There may be combinations of cluster offtake splits that result in more extreme behavior. The univariate analysis could result in one developing an overoptimistic view of how well one understands the model response over the entire parameter space.

Some Machine Learning algorithms, such as the Random Forest method, have built-in functionality that allows one to capture the relative importance of features in the multivariate context. One such technique, involves systematically switching the information for a feature between different samples, to remove the information that it contains, Reference [26]. A Random Forest model is first built for the prediction target of choice, using the training data, with performance metrics assessed, out-of-sample using test data (i.e. performing a blind test). For each of the features in turn, the values are permuted and re-run through the trained Random Forest model. For some of the features, there will be a detrimental impact on the accuracy of the predictions, as measured by the mean square prediction error (mse). If the feature in question is important, one would expect a large impact on the performance. If a relatively small impact on the performance is observed, then it can be inferred that the feature is not particularly important. The analysis can be repeated for each of the features in turn.

Whether the estimated importance in terms of the permutation based increase in mse of a feature is statistically significant can be determined by applying a variant of Student's t-test. The procedure is implemented for random forests in the Boruta algorithm and the associated R package. Further details can be found in Reference [28].

F.5 Partial Dependence Analysis

The relative variable importance analysis does not indicate whether an increase in the feature in question would cause an increase or decrease to the prediction target considered. For simple methods like a linear regression model, it is easy to understand the underlying relationship between the features and the prediction target because one can look at the sign and magnitude of the model coefficients. A simple parametric description cannot be examined in a similar fashion for the Random Forest.

Instead we can use the concept of a partial dependence plot, introduced by Friedman in Reference [29], that can be used for any so-called "black-box" style method. A Random Forest model is first built using the training data. The partial dependence analysis is then carried out, based on this model. Note that the partial dependence can be estimated for any subset of variables, such that interaction effects can be better understood. Let C denote the set of n controls on which the model response depends. We consider a subset C_S of C, such that $S \subset \{1, \dots, n\}$, and the complement set of controls is defined as $C_{\bar{S}}$, such that $\bar{S} = \{1, \dots, n\} \setminus S$. For a set of controls C, the response of the Random Forest model is given by:

$$f(C) = f(C_S, C_{\bar{S}}).$$

The partial dependence of the model on the controls specified in S is then defined as:

$$\overline{f}_s(C_S) = \frac{1}{p}\sum_{i=1}^{p} f(C_S, c_{i\bar{S}}),$$

where the $c_{i\bar{S}}$ are the values of the controls $C_{\bar{S}}$ of the i-th point of the training data on which $f(C)$ was trained.

In order to graph the relationship, we evaluate the partial dependence function $\overline{f}_s(C_S)$ on a regular grid spanned between the minimally and maximally observed values of C_s in the training set. Compare Reference [25] for further details. The procedure can be performed for each of the subsets of the features, in turn. This can be a computationally expensive process but for decision tree based methods, the partial dependence can be rapidly computed, directly from the trees, Reference [25].

Note that this partial dependence analysis provides more than merely the marginal dependence between the feature and prediction target. It represents the average effect of the feature in question on the predictions of the Random Forest model after controlling for the effects of all other features on the Random Forest model.

In this study, we use the partial dependence analysis primarily for a qualitative indication of whether, for a fixed production volume, increasing or decreasing the relative offtake percentage of a cluster would be beneficial for reducing the prediction target (such as population weighted PGV, max PGV etc.). We found that the overall relationships are mostly linear and that there are no strong interaction effects (section 6.7). If that were not the case, higher dimensional partial dependence plots would be required. Now since the effects are mostly linear and without strong interactions, we can illustrate the qualitative impact of changes in cluster controls by colouring the variable importance plot appropriately, with red indicating that the relative offtake split should be decreased in order to reduce the prediction target, green indicating that the relative offtake split should be increased and grey used for situations in which there is no statistically significant relationship between a control and the model response. Several illustrative example plots for different prediction targets on branch 6 are contained in Figure F-4, Figure F-5, Figure F-6, and Figure F-7.

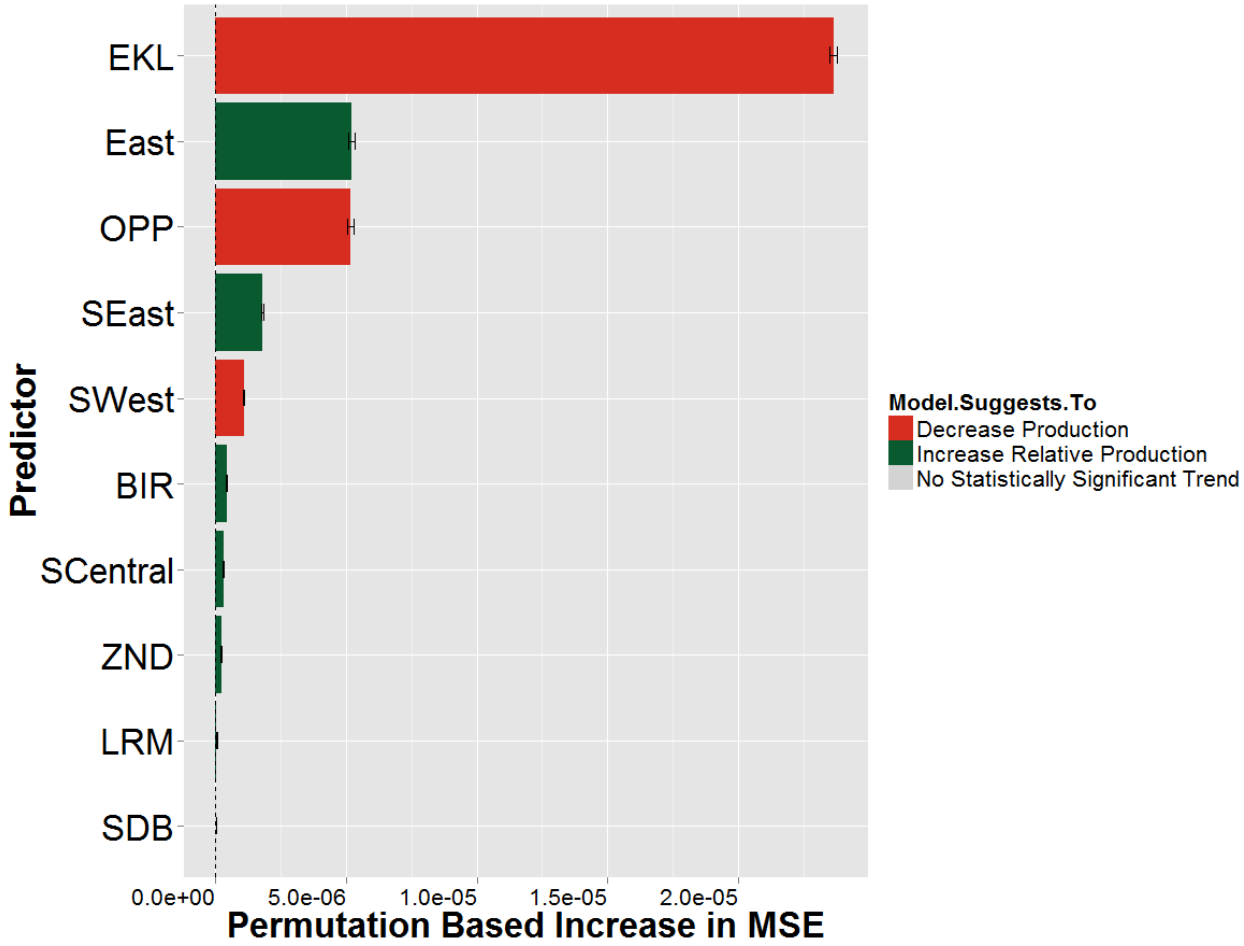


Figure F-4: Population weighted PGV for Branch Number 6, with 2365 samples that reach the production constraint

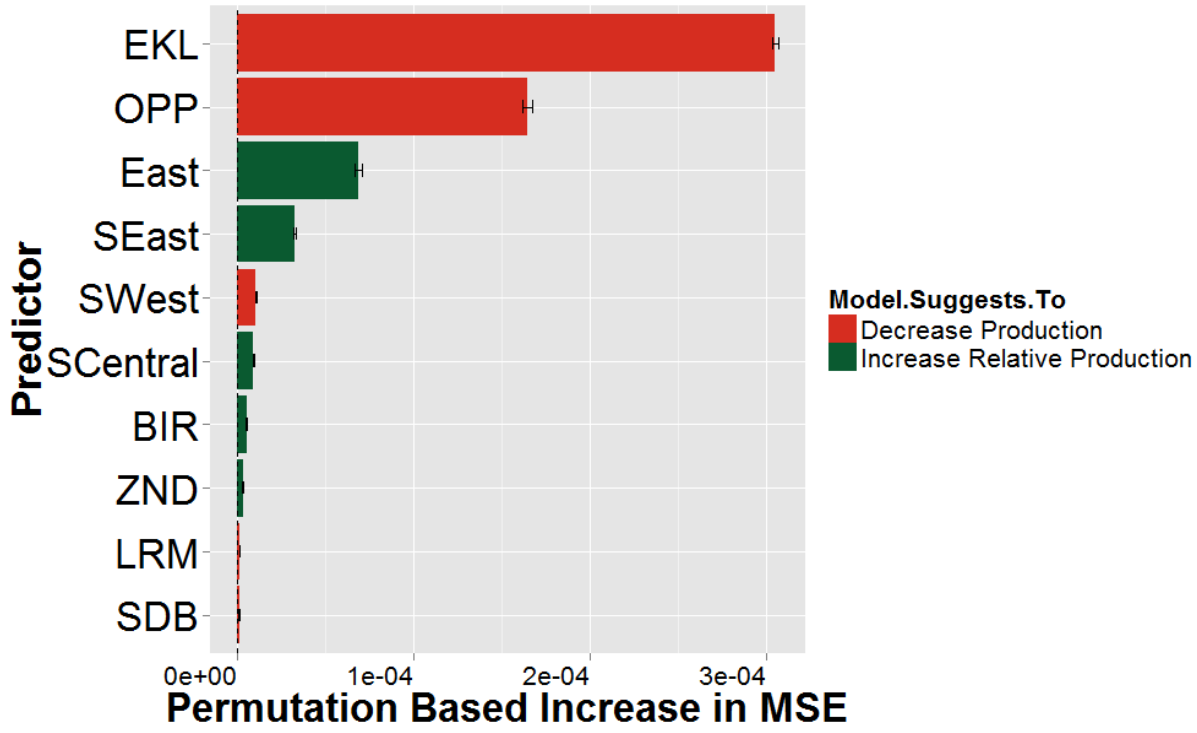


Figure F-5: Max PGA for Branch Number 6, with 2365 samples that reach the production constraint

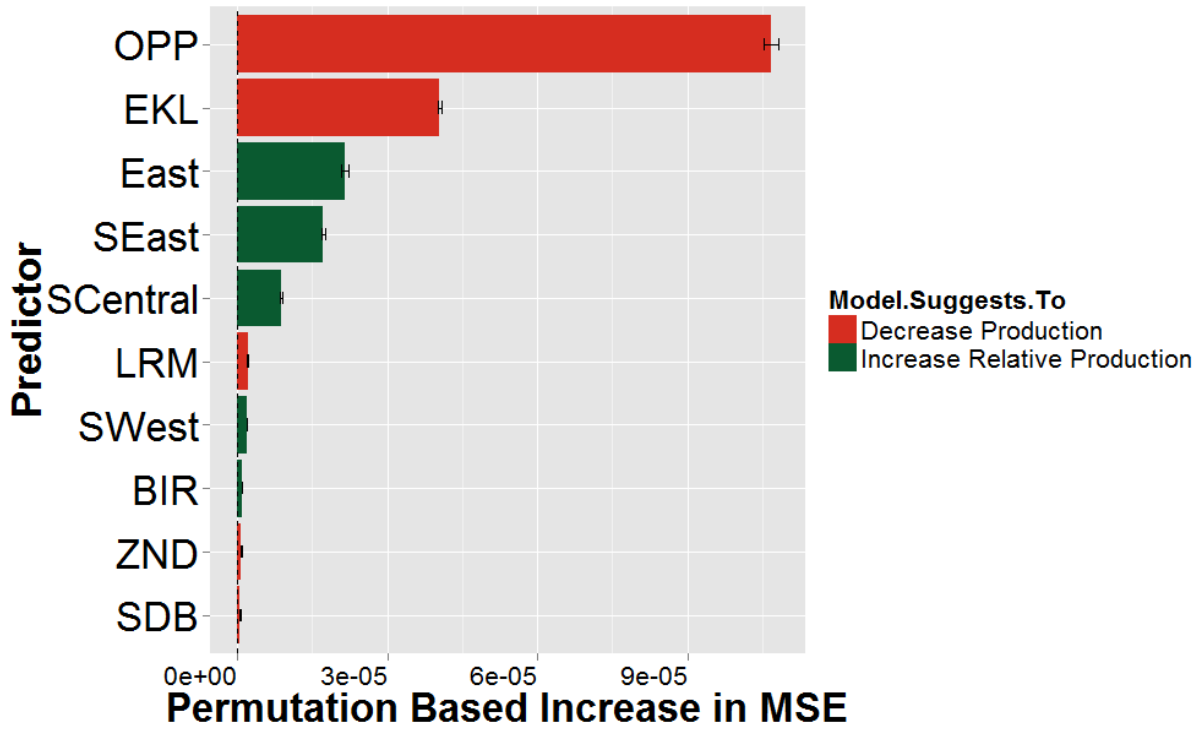


Figure F-6: Max PGV for Branch Number 6, with 2365 samples that reach the production constraint

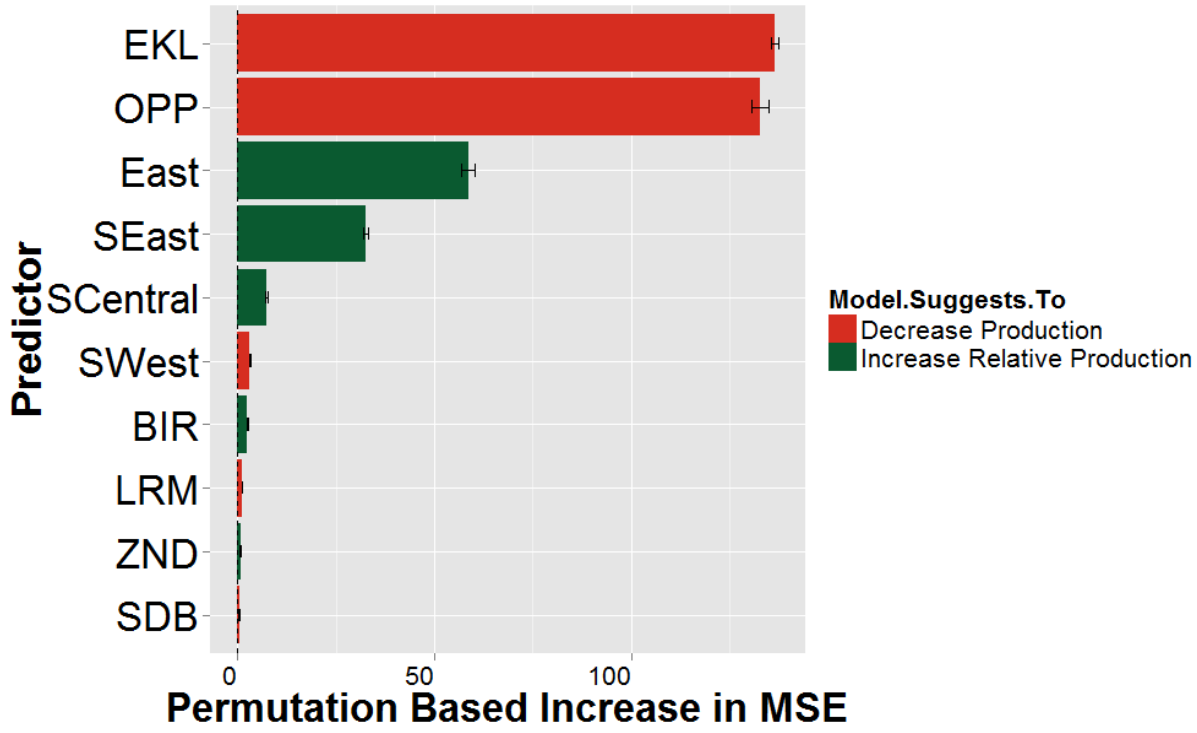


Figure F-7: Mean Number of Tremors, with 2365 samples that reach the production constraint

Appendix G List of Abbreviations

Bcm	N.Bcm refers to a volume of a billion normal cubic meters. Normal means the volume is measured at a standard temperature (0 degree C) and pressure (1 bar)
BHP	Bottom Hole Pressure
BP17	Business Plan 2017
DCS	Distributed Control System
EZ	Ministerie van Economische Zaken
GPS	Global Positioning System
Gron	Groningen
GEM	Global Earthquake Model (Global Science Forum of the OECD)
GenREM	Generalized Reservoir Evaluation Model
GMPE	Ground Motion Prediction Equation
GMM	Ground Motion Model
GPS	Global Positioning System
GTS	Gasunie Transport Services B.V.
GWC	Gas water contact
HRA	Hazard and Risk Assessment
InSAR	Interferometric Synthetic Aperture Radar
IPSM	Integrated Production System Model
JT valve	Joules Thompson valve
KNMI	Koninklijk Nederlands Meteorologisch Instituut
KO drum	Knock Out Drum
M	Earthquake Magnitude
ML	Local Earthquake Magnitude
M/Mw	Moment magnitude
MEA	Minister of Economic Affairs
Mmax	Maximum Earthquake Magnitude
MoReS	Modular Reservoir Simulator
MSE	Mean Square prediction Error
NAM	Nederlandse Aardolie Maatschappij B.V.
LTS	Low Temperature Separation
OV	Overslagen (gas custody transfer stations)
PGA	Peak Ground Acceleration
PGV	Peak Ground Velocity
pwPGV	Population Weighted PGV
SAC	Scientific Advisory Committee (Winningsplan 2016)
SodM	Staatstoezicht op de Mijnen (also SSM State Supervision of Mines)
SPMI	Simultaneous Perturbation and Multivariate Interpolation
THP	Tubing Head Pressure
TNO	Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek,

	Netherlands Organisation for Applied Scientific Research
UGS	Underground Gas Storage