

Szenariogenerierung zur Modellierung der stochastischen Ausspeiselasten in einem Gastransportnetz

Werner Römisch

(Dr. T. Koch (ZIB), H. Leövey, Dr. R. Mirkov, I. Wegner-Specht)

Institut für Mathematik
Humboldt-Universität Berlin



9. VDI-Fachtagung "Optimierung in der Energiewirtschaft"
Nürtingen 22./23.11.2011

Inhalt

Inhalt

- ▶ Gastransportnetz und Aufgabenstellung

Inhalt

- ▶ Gastransportnetz und Aufgabenstellung
- ▶ Statistische Daten und Datenanalyse

Inhalt

- ▶ Gastransportnetz und Aufgabenstellung
- ▶ Statistische Daten und Datenanalyse
- ▶ Schätzung der Wahrscheinlichkeitsverteilungen

Inhalt

- ▶ Gastransportnetz und Aufgabenstellung
- ▶ Statistische Daten und Datenanalyse
- ▶ Schätzung der Wahrscheinlichkeitsverteilungen
- ▶ Vorhersage der Gasabnahme im Tiefst-Temperaturbereich

Inhalt

- ▶ Gastransportnetz und Aufgabenstellung
- ▶ Statistische Daten und Datenanalyse
- ▶ Schätzung der Wahrscheinlichkeitsverteilungen
- ▶ Vorhersage der Gasabnahme im Tiefst-Temperaturbereich
- ▶ Generierung von Szenarien

Inhalt

- ▶ Gastransportnetz und Aufgabenstellung
- ▶ Statistische Daten und Datenanalyse
- ▶ Schätzung der Wahrscheinlichkeitsverteilungen
- ▶ Vorhersage der Gasabnahme im Tiefst-Temperaturbereich
- ▶ Generierung von Szenarien
- ▶ Szenarioerduktion

Gastransportnetz und Aufgabenstellung

Gastransportnetz und Aufgabenstellung

- ▶ Transportnetze von [Open Grid Europe GmbH \(OGE\)](#) für H- und L-Gas (hier: H-Gas),

Gastransportnetz und Aufgabenstellung

- ▶ Transportnetze von **Open Grid Europe GmbH (OGE)** für H- und L-Gas (hier: H-Gas),
- ▶ Aufgabenstellung: **Optimierung der technischen Kapazität von Gasnetzen,**

Gastransportnetz und Aufgabenstellung

- ▶ Transportnetze von **Open Grid Europe GmbH (OGE)** für H- und L-Gas (hier: H-Gas),
- ▶ Aufgabenstellung: **Optimierung der technischen Kapazität von Gasnetzen**,
- ▶ Speziell: **Buchungvalidierung**,

Gastransportnetz und Aufgabenstellung

- ▶ Transportnetze von **Open Grid Europe GmbH (OGE)** für H- und L-Gas (hier: H-Gas),
- ▶ Aufgabenstellung: **Optimierung der technischen Kapazität von Gasnetzen**,
- ▶ Speziell: **Buchungsvalidierung**,
- ▶ Spezifika: **hochdimensional**, **stochastisch** nichtlinear, gemischt-ganzzahlig.

Statistische Daten

Statistische Daten

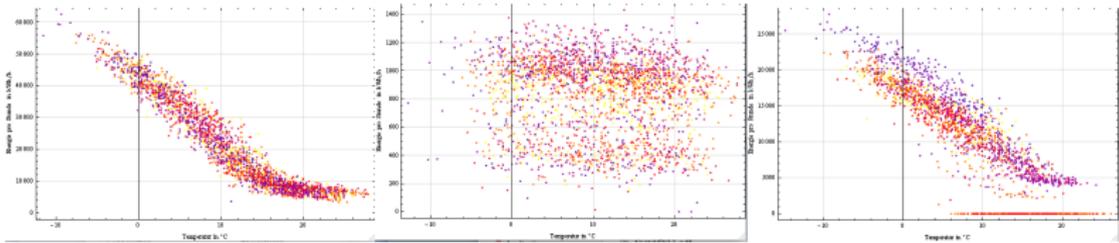
- ▶ Datenbasis: historische Daten für Gasflussmengen der Abnehmer an Auspeisepunkten über einem längeren Zeitraum,

Statistische Daten

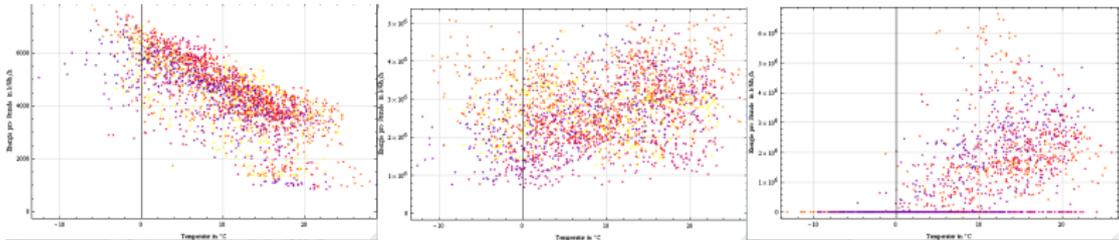
- ▶ Datenbasis: historische Daten für Gasflussmengen der Abnehmer an Ausseispunkten über einem längeren Zeitraum,
- ▶ wesentlich sind eine ausreichende Anzahl von Messungen und die Vollständigkeit der Daten,

Statistische Daten

- ▶ Datenbasis: historische Daten für Gasflussmengen der Abnehmer an Ausspeisepunkten über einem längeren Zeitraum,
- ▶ wesentlich sind eine ausreichende Anzahl von Messungen und die Vollständigkeit der Daten,
- ▶ letzteres hängt wesentlich vom Verwendungszweck ab (z.B. für dynamische oder stationäre Modellierung).



Datensätze von Ausspeiseknoten mit Stadtwerken, rechts auch mit Nullausspeisung.



Datensätze von einem Knoten mit Industrieannehmer (links), Marktübergang (Mitte), Speicher (rechts).

Datenanalyse

Datenanalyse

- ▶ **Stationäre Modellierung:** Sicherung der Vergleichbarkeit von Daten, Daten sind temperaturabhängig.

Datenanalyse

- ▶ **Stationäre Modellierung:** Sicherung der Vergleichbarkeit von Daten, Daten sind temperaturabhängig.
- ▶ Beispiel für stationäre Modellierung:

Datenanalyse

- ▶ **Stationäre Modellierung:** Sicherung der Vergleichbarkeit von Daten, Daten sind temperaturabhängig.
- ▶ Beispiel für stationäre Modellierung:
 - ▶ Verwendung des mittleren Gasflusses,

Datenanalyse

- ▶ **Stationäre Modellierung:** Sicherung der Vergleichbarkeit von Daten, Daten sind temperaturabhängig.
- ▶ Beispiel für stationäre Modellierung:
 - ▶ Verwendung des mittleren Gasflusses,
 - ▶ Verwendung einer räumlich gemittelten Referenztemperatur,

Datenanalyse

- ▶ **Stationäre Modellierung:** Sicherung der Vergleichbarkeit von Daten, Daten sind temperaturabhängig.
- ▶ Beispiel für stationäre Modellierung:
 - ▶ Verwendung des mittleren Gasflusses,
 - ▶ Verwendung einer räumlich gemittelten Referenztemperatur,
 - ▶ Bestimmung von Temperaturklassen: $(-15,-4]$, $(-4,-2]$, $(-2,0]$, \dots , $(18,20]$, $(20,40)$.

Datenanalyse

- ▶ **Stationäre Modellierung:** Sicherung der Vergleichbarkeit von Daten, Daten sind temperaturabhängig.
- ▶ Beispiel für stationäre Modellierung:
 - ▶ Verwendung des mittleren Gasflusses,
 - ▶ Verwendung einer räumlich gemittelten Referenztemperatur,
 - ▶ Bestimmung von Temperaturklassen: $(-15,-4]$, $(-4,-2]$, $(-2,0]$, \dots , $(18,20]$, $(20,40)$.
 - ▶ entsprechende Filterung der Abnehmerdaten.

Temperaturklassen

Temperaturklassen

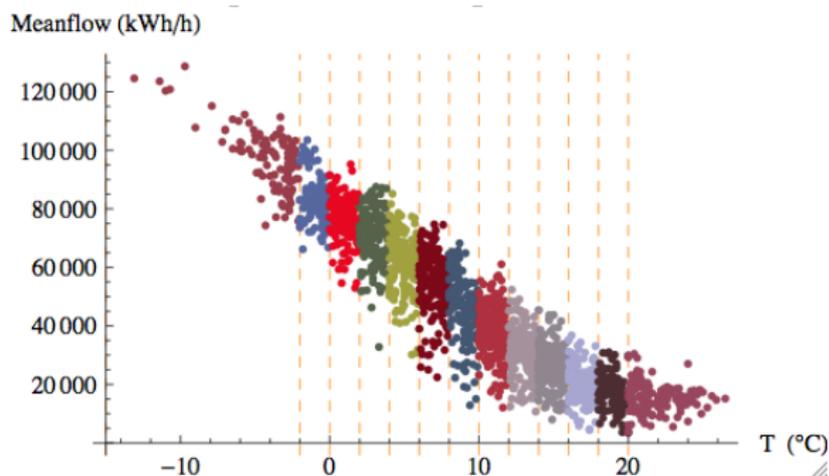
- ▶ Für jeden Ausspeiseknoten und jede Temperaturklasse werden alle Tage und die zugehörigen mittleren Gasflüsse gefiltert,

Temperaturklassen

- ▶ Für jeden Ausspeiseknoten und jede Temperaturklasse werden alle Tage und die zugehörigen mittleren Gasflüsse gefiltert,
- ▶ Daten-Preprocessing zur Überprüfung der statistischen Belastbarkeit eines Exit-Knotens in allen T-Klassen.

Temperaturklassen

- ▶ Für jeden Ausspeiseknoten und jede Temperaturklasse werden alle Tage und die zugehörigen mittleren Gasflüsse gefiltert,
- ▶ Daten-Preprocessing zur Überprüfung der statistischen Belastbarkeit eines Exit-Knotens in allen T-Klassen.



Anpassung von univariaten Verteilungen

Anpassung von univariaten Verteilungen

Klassen von Verteilungen

- ▶ (shifted) Gleichverteilungen

Anpassung von univariaten Verteilungen

Klassen von Verteilungen

- ▶ (shifted) Gleichverteilungen
- ▶ (shifted) Normalverteilungen

Anpassung von univariaten Verteilungen

Klassen von Verteilungen

- ▶ (shifted) Gleichverteilungen
- ▶ (shifted) Normalverteilungen
- ▶ bei zahlreichen Knoten tritt eine Null-Ausspeisung mit empirischer Häufigkeit p auf.

Anpassung von univariaten Verteilungen

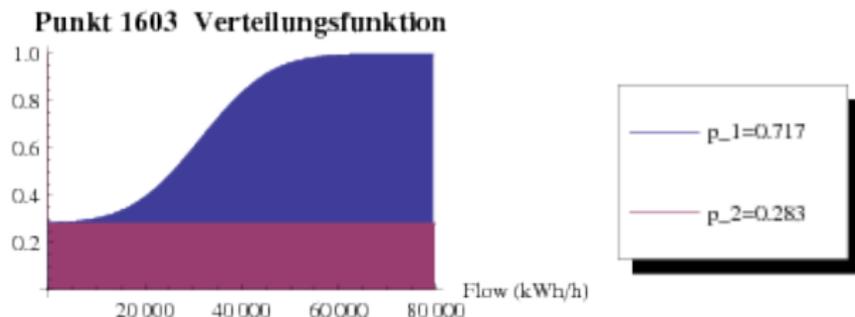
Klassen von Verteilungen

- ▶ (shifted) Gleichverteilungen
- ▶ (shifted) Normalverteilungen
- ▶ bei zahlreichen Knoten tritt eine Null-Ausspeisung mit empirischer Häufigkeit p auf.
- ▶ Ansatz für Verteilungsfunktion:
$$F(x) = p F^0(x) + (1 - p) F^+(x).$$

Anpassung von univariaten Verteilungen

Klassen von Verteilungen

- ▶ (shifted) Gleichverteilungen
- ▶ (shifted) Normalverteilungen
- ▶ bei zahlreichen Knoten tritt eine Null-Ausspeisung mit empirischer Häufigkeit p auf.
- ▶ Ansatz für Verteilungsfunktion:
$$F(x) = p F^0(x) + (1 - p) F^+(x).$$
- ▶ Verteilungsfunktion einer shifted Normalverteilung



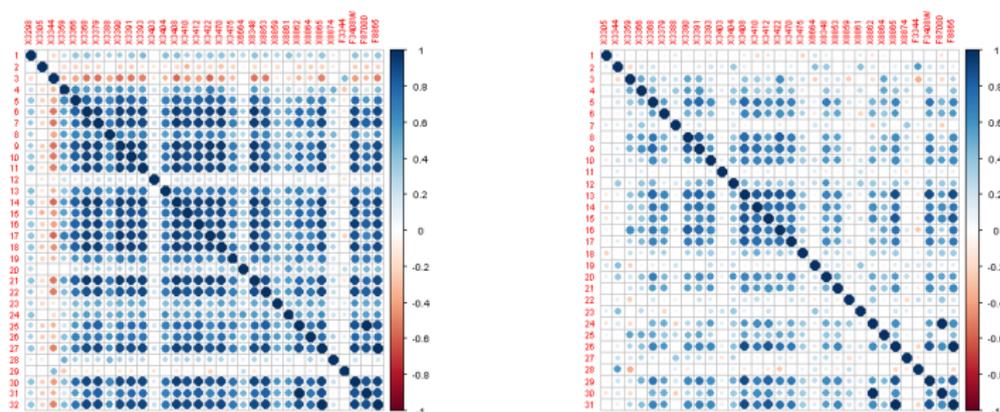
Anpassung multivariater Normalverteilungen

Anpassung multivariater Normalverteilungen

- ▶ Für Ausspeisepunkten, die Normalverteilungstest bestehen und signifikante Korrelationen zu anderen Knoten aufweisen, werden multivariate Normalverteilungen angepasst, d.h. zusätzlich zu Mittelwert und Varianz werden Korrelationen geschätzt (Standardschätzer nach Pearson).

Anpassung multivariater Normalverteilungen

- ▶ Für Ausseispunkten, die Normalverteilungstest bestehen und signifikante Korrelationen zu anderen Knoten aufweisen, werden multivariate Normalverteilungen angepasst, d.h. zusätzlich zu Mittelwert und Varianz werden Korrelationen geschätzt (Standardschätzer nach Pearson).
- ▶ Beispiele für Korrelationsmatrizen:



Korrelationsplots für die Temperaturklassen (10,12] und (18,20].

Vorhersage der Gasabnahme im Tief-T-Bereich

Vorhersage der Gasabnahme im Tief-T-Bereich

- ▶ Spezielle Behandlung, da oftmals nicht genügend Daten vorhanden sind.

Vorhersage der Gasabnahme im Tief-T-Bereich

- ▶ Spezielle Behandlung, da oftmals nicht genügend Daten vorhanden sind.
- ▶ **Idee:** P-Spline Regression mit Form-Restriktionen.

Vorhersage der Gasabnahme im Tief-T-Bereich

- ▶ Spezielle Behandlung, da oftmals nicht genügend Daten vorhanden sind.
- ▶ **Idee:** P-Spline Regression mit Form-Restriktionen.
- ▶ **Ansatz:** Methode der kleinsten Quadrate für alle y_i (standardisierte maximale tägliche Gasflusswerte) zur Temperatur t_i eines Ausspeisepunktes

$$\sum_{i=1}^n (y_i - S(t_i))^2 + \lambda \sum_{j=3}^m (\delta^2 a_j)^2 + \kappa \sum_{j=2}^m b_j (\delta^1 a_j)^2,$$

wobei λ und κ Glättungs- bzw. Form-Parameter sind und

$$S(t) = \sum_{j=1}^m a_j B_j(t)$$

ein kubischer Spline in B-Spline Darstellung ist. Hierbei ist $\delta^1 a_j = a_j - a_{j-1}$ und $\delta^2 a_j = a_j - 2a_{j-1} + a_{j-2}$ gilt.

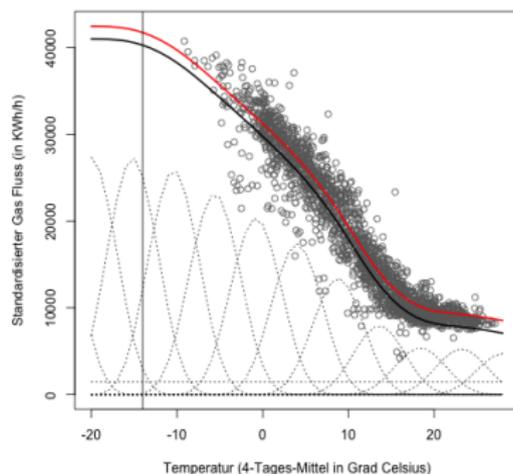
Vorhersage der Gasabnahme im Tief-T-Bereich 2

Vorhersage der Gasabnahme im Tief-T-Bereich 2

- ▶ Anpassung einer univariaten Normalverteilung mit Mittelwert aus dem Regressionsmodell bei $t = -14^{\circ}\text{C}$ und Standardabweichung der Temperaturklasse $(-4^{\circ}\text{C}, -2^{\circ}\text{C}]$.

Vorhersage der Gasabnahme im Tief-T-Bereich 2

- ▶ Anpassung einer univariaten Normalverteilung mit Mittelwert aus dem Regressionsmodell bei $t = -14^{\circ}\text{C}$ und Standardabweichung der Temperaturklasse $(-4^{\circ}\text{C}, -2^{\circ}\text{C}]$.
- ▶ Beispiel Parameter: $\lambda = 2.51$, $\kappa = 100$ und $b_1 = b_2 = 1$.



P-Spline Regression mit abflachender Asymptote.

Generierung von Szenarien

Generierung von Szenarien

- ▶ Verwendung von randomisierten Quasi-Monte Carlo Methoden,

Generierung von Szenarien

- ▶ Verwendung von **randomisierten Quasi-Monte Carlo Methoden**,
- ▶ Für einen d -dimensionalen Zufallsvektor, der allen statistischen Exits des Transportnetzes entspricht, werden N mögliche Realisierungen mit den Wahrscheinlichkeiten $\frac{1}{N}$ erzeugt,

Generierung von Szenarien

- ▶ Verwendung von **randomisierten Quasi-Monte Carlo Methoden**,
- ▶ Für einen d -dimensionalen Zufallsvektor, der allen statistischen Exits des Transportnetzes entspricht, werden N mögliche Realisierungen mit den Wahrscheinlichkeiten $\frac{1}{N}$ erzeugt,
- ▶ **Schritte:**

Generierung von Szenarien

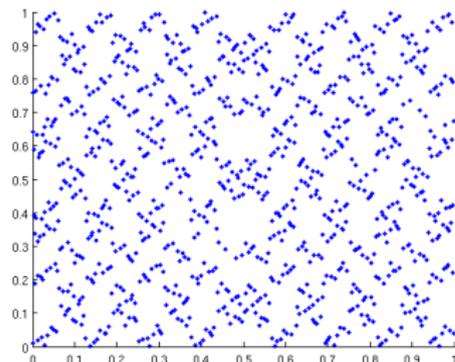
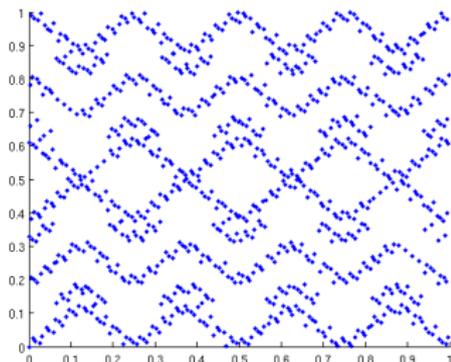
- ▶ Verwendung von **randomisierten Quasi-Monte Carlo Methoden**,
- ▶ Für einen d -dimensionalen Zufallsvektor, der allen statistischen Exits des Transportnetzes entspricht, werden N mögliche Realisierungen mit den Wahrscheinlichkeiten $\frac{1}{N}$ erzeugt,
- ▶ **Schritte:**
 - ▶ Bestimmung von N möglichen Realisierungen η^j einer Gleichverteilung auf $[0, 1)^d$ aus der sog. Niederreiter-Folge zur Basis 2.

Generierung von Szenarien

- ▶ Verwendung von **randomisierten Quasi-Monte Carlo Methoden**,
- ▶ Für einen d -dimensionalen Zufallsvektor, der allen statistischen Exits des Transportnetzes entspricht, werden N mögliche Realisierungen mit den Wahrscheinlichkeiten $\frac{1}{N}$ erzeugt,
- ▶ **Schritte:**
 - ▶ Bestimmung von N möglichen Realisierungen η^j einer Gleichverteilung auf $[0, 1)^d$ aus der sog. Niederreiter-Folge zur Basis 2.
 - ▶ **Random Scrambling** der Folge nach Owen, d.h. zufällige Permutation der Dezimalziffern jeder Komponente (Mersenne-Twister).

Generierung von Szenarien 2

Generierung von Szenarien 2



Links: 1 000 Niederreiter-Punkte für $d = 40$, Projektion (16, 18).

Rechts: 1 000 Scrambled-Niederreiter-Punkte für $d = 40$, Projektion (16, 18).

Generierung von Szenarien 3

Generierung von Szenarien 3

- ▶ Die Szenarien η^j , $i = 1, \dots, N$, vereinen positive Eigenschaften von QMC und MC Methoden.

Generierung von Szenarien 3

- ▶ Die Szenarien η^j , $i = 1, \dots, N$, vereinen positive Eigenschaften von QMC und MC Methoden.
- ▶ Bestimmung neuer Vektoren

$$\zeta_i^j = \Phi_i^{-1}(\eta_i^j) \quad (i = 1, \dots, d; j = 1, \dots, N)$$

mit Hilfe der (univariaten) Verteilungsfunktionen der i ten Komponente. Im Fall der Normalverteilung wird die Standardnormalverteilung $\mathcal{N}(0, 1)$ verwendet.

Generierung von Szenarien 3

- ▶ Die Szenarien η^j , $i = 1, \dots, N$, vereinen positive Eigenschaften von QMC und MC Methoden.
- ▶ Bestimmung neuer Vektoren

$$\zeta_i^j = \Phi_i^{-1}(\eta_i^j) \quad (i = 1, \dots, d; j = 1, \dots, N)$$

mit Hilfe der (univariaten) Verteilungsfunktionen der i ten Komponente. Im Fall der Normalverteilung wird die Standardnormalverteilung $\mathcal{N}(0, 1)$ verwendet.

- ▶ Hat man eine s -dimensionale multivariate Normalverteilung $\mathcal{N}(m, \Sigma)$ mit Mittelwertsvektor $m \in \mathbb{R}^s$ und $s \times s$ -dimensionaler Kovarianzmatrix Σ , so bestimmt man die sog. Cholesky-Zerlegung $\Sigma = L L^\top$ der Kovarianzmatrix Σ mit einer unteren Dreiecksmatrix L vom Typ $s \times s$. Dann sind die s -dimensionalen Vektoren

$$\xi^j = L \zeta^j + m \quad (j = 1, 2, \dots, N)$$

geeignete Szenarien für $\mathcal{N}(m, \Sigma)$.

Szenarioreduktion

Szenarioreduktion

- ▶ Gegeben sei eine große Anzahl N von d -dimensionalen Szenarien ξ^i , $i = 1, \dots, N$, mit Wahrscheinlichkeiten $\frac{1}{N}$.

Szenarioreduktion

- ▶ Gegeben sei eine große Anzahl N von d -dimensionalen Szenarien ξ^i , $i = 1, \dots, N$, mit Wahrscheinlichkeiten $\frac{1}{N}$.
- ▶ Optional sollen nun $n < N$ Szenarien daraus ausgewählt und die zu diesen ausgewählten Szenarien gehörigen **neuen Wahrscheinlichkeiten** bestimmt werden, so dass das große Ensemble von N Szenarien **bestmöglich approximiert** wird.

Szenarioreduktion

- ▶ Gegeben sei eine große Anzahl N von d -dimensionalen Szenarien ξ^i , $i = 1, \dots, N$, mit Wahrscheinlichkeiten $\frac{1}{N}$.
- ▶ Optional sollen nun $n < N$ Szenarien daraus ausgewählt und die zu diesen ausgewählten Szenarien gehörigen **neuen Wahrscheinlichkeiten** bestimmt werden, so dass das große Ensemble von N Szenarien **bestmöglich approximiert** wird.
- ▶ Dies erfordert die Lösung eines kombinatorischen Optimierungsproblems. Dazu werden Heuristiken eingesetzt.

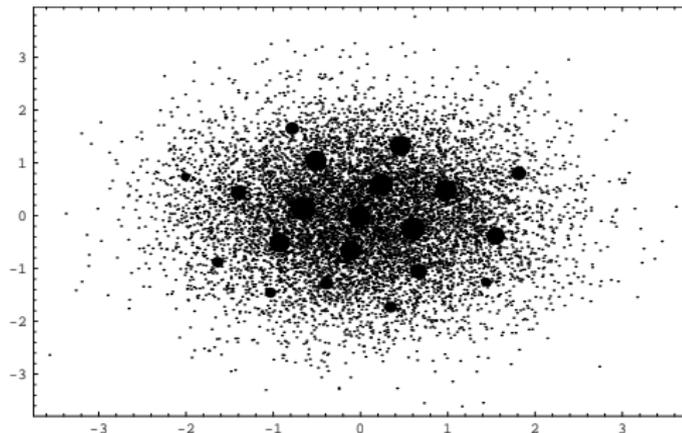
Szenarioreduktion 2

Szenarioreduktion 2

- ▶ Die Wahrscheinlichkeit jedes gestrichenen Szenarios wird zu der eines nächstgelegenen verbleibenden Szenarios addiert.

Szenarioreduktion 2

- ▶ Die Wahrscheinlichkeit jedes gestrichenen Szenarios wird zu der eines nächstgelegenen verbleibenden Szenarios addiert.
- ▶ **Illustration:** Ausgangspunkt sind $N = 10\,000$ erzeugte Szenarien der zweidimensionalen Normalverteilung. Reduktion auf $n = 20$. Die Durchmesser der Kreise sind proportional zu ihren neuen Wahrscheinlichkeiten.



Zusammenfassung

Zusammenfassung

- ▶ Durchführung einer deskriptiven Datenanalyse und Filterung der mittleren täglichen Ausspeiselasten aller Exits in Temperaturklassen.

Zusammenfassung

- ▶ Durchführung einer deskriptiven Datenanalyse und Filterung der mittleren täglichen Ausspeiselasten aller Exits in Temperaturklassen.
- ▶ Es werden univariate (shifted) Gleich- und Normalverteilungen sowie multivariate Normalverteilungen für (Gruppen von) Exits geschätzt.

Zusammenfassung

- ▶ Durchführung einer deskriptiven Datenanalyse und Filterung der mittleren täglichen Ausspeiselasten aller Exits in Temperaturklassen.
- ▶ Es werden univariate (shifted) Gleich- und Normalverteilungen sowie multivariate Normalverteilungen für (Gruppen von) Exits geschätzt.
- ▶ Für den Tiefsttemperaturbereich mit zu wenig Daten wird eine univariate Normalverteilung angenommen, deren Mittelwert mit Hilfe eines P-Spline Regressionsmodells bestimmt wird.

Zusammenfassung

- ▶ Durchführung einer deskriptiven Datenanalyse und Filterung der mittleren täglichen Ausspeiselasten aller Exits in Temperaturklassen.
- ▶ Es werden univariate (shifted) Gleich- und Normalverteilungen sowie multivariate Normalverteilungen für (Gruppen von) Exits geschätzt.
- ▶ Für den Tiefsttemperaturbereich mit zu wenig Daten wird eine univariate Normalverteilung angenommen, deren Mittelwert mit Hilfe eines P-Spline Regressionsmodells bestimmt wird.
- ▶ Mittels randomisierter Quasi-Monte Carlo methoden wird eine größere Anzahl von Szenarien der Exits erzeugt.

Zusammenfassung

- ▶ Durchführung einer deskriptiven Datenanalyse und Filterung der mittleren täglichen Ausspeiselasten aller Exits in Temperaturklassen.
- ▶ Es werden univariate (shifted) Gleich- und Normalverteilungen sowie multivariate Normalverteilungen für (Gruppen von) Exits geschätzt.
- ▶ Für den Tiefsttemperaturbereich mit zu wenig Daten wird eine univariate Normalverteilung angenommen, deren Mittelwert mit Hilfe eines P-Spline Regressionsmodells bestimmt wird.
- ▶ Mittels randomisierter Quasi-Monte Carlo methoden wird eine größere Anzahl von Szenarien der Exits erzeugt.
- ▶ Optional kann durch Szenarioerduktion eine geringere Anzahl von Szenarien und deren Wahrscheinlichkeiten bestimmt werden, die das große Ensemble von ursprünglichen Szenarien bestmöglich repräsentieren.