

ベイズ・ニューラルネットワーク法による 土壌汚染浄化費用の推定

平成18年2月20日

京都大学工学部地球工学科土木コース

徐 飛

要 旨

本研究では、表土調査の結果に基づいて、土壤汚染浄化費用リスクを計測する方法論を提案する。表土調査の結果をインプットとし、土壤汚染費用をアウトプットとするニューラルネットワークモデルを作成する。伝統的なニューラルネットワークモデルでは、アウトプットを点推定を行うことを目的とする。これに対して、本研究では、表土調査の結果に基づいて推計する土壤汚染費用の推計結果には推計誤差リスクが存在することに着目する。本研究では、ベイズ推計法によりニューラルネットワークモデルを推計することにより、土壤汚染処理費用の確率分布を推計する方法論を提案することを目的とする。さらに、新たに土壤汚染サンプルを獲得することにより、ベイズニューラルネットワークモデルを更新する方法論も提案する。さらに、現実の土壤汚染費用データベースに基づいて、提案したベイズニューラルネットワークモデルの有効性について実証的に検証する。

目 次

1 はじめに	1
2 本研究の基本的な考え方	2
2.1 従来の研究概要	2
2.2 土壌汚染の調査方法	3
2.3 土壌汚染浄化費用リスク	4
2.4 土壌汚染浄化費用リスクの評価プロセス	5
3 BNNモデルの定式化	6
3.1 ベイズ予測問題	6
3.2 NNモデル（プロセス1）	7
3.3 BNNモデル（プロセス2）	9
4 土壌汚染浄化費用リスクの評価	12
4.1 リスク評価の方法	12
4.2 MHアルゴリズム	12
4.3 事後分布に関する統計量	15
4.4 土壌汚染浄化費用リスクの評価	18
4.5 BNNモデルのベイズ更新	19
5 適用事例	21
5.1 適用事例の概略	21
5.2 BNNモデルの推計	21
5.3 土壌汚染浄化費用リスクのベイズ予測	23
6 おわりに	24

1 はじめに

近年、土地資産の売却および担保評価、不動産取引、あるいは企業の工場跡地等の再開発にともない、市街地における土壤汚染問題が顕在化してきた。土壤汚染が原因となって生じる健康リスク、環境リスクは無視できない。さらに、地下水汚染等が原因となり、土壤汚染被害が広範囲に及ぶ可能性がある。この場合、土壤汚染が第3者に多大な外部不経済をもたらすことになる。このような状況を背景として、土壤汚染に関する法律「土壤汚染対策法」が2003年3月に施行された¹⁾。

同法の施行により、土地所有者が土地開発を実施する際、土壤汚染の実態を調査することが義務付けられた。しかし、土壤汚染浄化費用は高額であり、案件によっては土壤汚染浄化費用が数億円に達し、事業自体が断念された例もある²⁾。このため、土地開発にあたり、土壤汚染浄化費用が重要なリスク管理項目となった。土壤汚染の実態を把握するためには、表土調査、ボーリング調査等による現地調査が必要となる。しかし、ボーリング調査を行うためには相当程度の調査費用が必要となる。さらに、調査範囲が大きくなれば、事業者にとって調査費用自体が大きな負担となる。

土地開発の採算性や実現可能性を検討するような構想段階においては、多額の調査費用を費やし、詳細な土壤汚染調査を実施することは合理的ではない。むしろ、地籍調査、表土調査等の簡易調査に基づいて土壤汚染浄化費用リスクを評価し、マーケットリスク等を含めて土地開発の妥当性を検討することが必要となる。しかし、地籍調査、表土調査等の簡易調査のみが実施された段階で、土壤汚染の実態を確定的に把握することは極めて困難である。したがって、土壤汚染浄化費用のありうべき確率分布の程度(以下、土壤汚染浄化費用リスクと呼ぶ)を評価することが重要となる。

以上の問題意識の下に、本研究では地籍調査、表土調査等の簡易調査結果に基づいて、土壤汚染浄化費用リスクを評価するためのベイズニューラルネットワーク(Bayesian Neural Network: 以下、BNNと略す)モデルを提案する。以下、2.において、本研究の基本的な考え方を説明する。3.で、土壤汚染浄化費用モデルを定式化する。4.でBNNモデルを用いて土壤汚染浄化費用リスクを評価するための方法論を提案する。5.で、BNNモデルの適用事例について考察する。

2 本研究の基本的な考え方

2.1 従来の研究概要

土壤汚染浄化費用リスクの評価に関しては、いくつかの研究事例がある。これらの研究事例は、土壤汚染物質の拡散メカニズムを考慮した力学的モデルと、土壤汚染浄化費用の実績値と地点特性の相関関係に基づいて土壤汚染浄化費用リスクを評価する統計的モデルに大別できる。前者の立場からは、ボーリング調査の結果に基づいて、土壤汚染の実態を地盤統計学や拡散モデル等を用いて推計する方法論が提案されている³⁾⁻⁵⁾。しかし、筆者等が知るかぎり、土地開発の構想段階を対象として、地籍調査、表土調査の結果から、力学的モデルを用いて土壤汚染浄化費用を推計しようとする実用的な試みは見いだせない。特に、表土調査からは、土壤汚染の深度に関する情報が得られないという限界がある。また、地盤条件が均一でない場合、表土調査に基づく土壤汚染浄化費用の推定には、多くの推計リスクが介在する。このため、簡易調査に基づいて土壤汚染浄化費用を見積もる場合、土壤汚染浄化費用リスクを明示的に評価できるような統計的モデルに頼らざるを得ない。このような観点から、本研究では、統計的モデルの1つであるBNNモデルを用いて土壤汚染浄化費用リスクを評価する方法論を提案することとする。

ニューラルネットワーク（以下、NNと略す）モデルに関しては、すでに膨大な研究蓄積⁶⁾⁻⁸⁾があり、土木工学の分野でも適用事例は数多い。NNモデルは、複雑な構造を有する非線形現象を容易にモデル化できる利点を持っている。特に、少なくとも1つの中間層を有する多層ニューラルネットでコンパクト集合上で定義された任意の関数を近似できることが証明された^{9),10)}。その結果、最近では、NNモデルを柔軟な構造を有する非線形予測モデルとして位置づけるようになってきた。一方、ベイズ統計学は、長い研究の歴史があり¹¹⁾⁻¹³⁾、土木工学の分野でもいくつかの適用事例がある¹⁴⁾。しかし、ベイズ統計学に基づいたNNモデルに関する研究は緒についたばかりで、研究事例は極めて少ないのが実情である。のちに、4.(1)で言及するように、NNモデルにベイズ推計法を適用する場合、モデルの重み係数の共役分布を解析的に求めることが不可能である。このことが、ベイズ推計法の実用化の障害になってきた。また、NNモデルが重み係数を学習するという立場に立っており、このことが同じく統計的学習という立場に立つベイズ統計学と相容れな

いという一般的な認識が存在していたと言われる¹⁵⁾。しかし、近年になりマルコフ連鎖モンテカルロ法 (Markov Chain Monte Carlo Method; 以下、MCMC法という) の発展により、容易に事後分布を求めることが可能になってきた¹⁶⁾。さらに、Radford は確率誤差項を導入したNNモデルを提案することにより、NNモデルの学習過程にベイズ統計学を適用したBNNモデルを提案した¹⁵⁾。本研究では、BNNモデルを用いて土壌汚染浄化費用をベイズ予測する方法論を提案する。これにより、土壌汚染に関する簡易調査の結果に基づいて、調査地点の土壌汚染浄化費用リスクを評価することが可能となる。筆者等の知る限り、ベイズ統計学を用いて土壌汚染浄化費用リスクを評価した事例は見あたらない。

2.2 土壌汚染の調査方法

土地所有者が土地開発を実施する際に、土壌汚染調査を実施することが義務づけられている。土地の汚染状況を把握するための調査としては、1) 地籍調査、2) 表土調査、及び3) ボーリング調査等の調査方法が存在する。当然のことながら、これら3つの調査方法は、それぞれ調査費用や調査結果の精度が異なっている。土地開発に関する企画・調査の検討段階やその内容に応じて、必要となる土壌汚染調査が異なってくる。

地籍調査は最も簡易的な調査手法である。地籍調査により、対象土地の用途地域、周辺状況、地勢、土地利用状況、敷地規模などの地点特性情報と、過去の時点における土地利用の履歴や土壌汚染事故の有無等の情報が得られる。土地開発の調査・企画の初期段階において、土壌汚染の有無に関する定性的な判断が求められる時に、地籍調査結果が用いられる。土地開発の企画がある程度具体的になり、いくつかの土地開発企画案の経営的妥当性やリスク評価を実施する段階に進めば、土壌汚染の有無だけでなく、(土壌汚染が存在する場合) 土壌汚染浄化費用を概算することが求められる。この場合、表土調査を実施し、土壌汚染浄化費用リスクを評価することが必要となる。表土調査は現地に立ち入って、汚染物の有無を直接調べる手法であり、同時に汚染の範囲、汚染物の種類及び地表での汚染物濃度も調べられる。しかし、その表土調査では土壌内の汚染状況を把握できず、土壌汚染浄化費用の推計結果には多くの計測誤差が含まれる。さらに、土地開発を実施することが具体化した場合、土壌汚染の浄化工法の選択や、より正確な土壌汚染浄化費用の見積もりが必要となる。この場合、ボーリング調査を実施し、土壌汚染の

実態をより正確に評価することが必要となる．ボーリング調査を実施すれば，地質状況，汚染の有無，範囲，深度，汚染物の種類，濃度，地下水位及び地下水汚染の有無などの詳細，精確かつ豊富な情報が得られる．ボーリング調査は，3つの調査方法の中で，土壤汚染の状況をもっとも的確に把握できるが，調査費用自体が相当な負担となる．

2.3 土壤汚染浄化費用リスク

「リスク」という用語は多義的であり，統一的な定義はない．本研究では「リスク」という用語を，確率分布という意味で用いる．土壤汚染は，現実存在しており，現象そのものにリスクが存在するわけではない．土壤汚染浄化対策を講じた事後の時点では，土壤汚染に関するリスクは存在しない．しかし，土壤汚染浄化対策を講じる事前の段階では，土壤汚染の実態を直接観測することは不可能である．その結果，土壤汚染浄化のために必要となる費用を確定的に把握できない．

2.(2)で言及したように，土壤汚染調査が地籍調査，表土調査，ボーリング調査と進展するにつれて，土壤汚染に関わるより詳細な情報を獲得することができる．それに伴って，リスクの程度は減少する．例えば，ボーリング調査を実施すれば，汚染物質の拡散モデルや地盤統計学等を駆使することにより，土壤汚染の実態をより詳細に把握できる．しかし，本研究では，前述したように，土地所有者が土地開発の企画・計画を実施する構想段階をとりあげる．この段階では，ボーリング調査等，多額の調査費用を支払って，土壤汚染の実態を調査することは合理的ではない．地籍調査，表土調査等の簡易な調査を実施し，土壤汚染浄化費用リスクを評価するとともに，マーケットリスク等を総合的に勘案し，土地開発の継続の有無や，土地開発の内容を決定するという高度なリスク判断が必要となる．

地籍調査，表土調査のみを行った段階では，地盤の不均一性に起因して，土壤汚染浄化費用の推計値の信頼性には問題が残る場合が少なくない．このような段階では，土壤汚染浄化費用を確定的に計測することは不可能であり，図－1に示すように，土壤汚染浄化費用の確率分布を測定することが現実的であろう．同図では，表土調査の結果に基づいて，土壤汚染浄化費用のありうべき確率分布を予測した結果を示している．本研究で提案するBNNモデルは，同図に示すような土壤汚染浄化費用リスクを求めることを目的とする．

2.4 土壤汚染浄化費用リスクの評価プロセス

NNモデルは構造的な自由度が大きく、試行錯誤を通じて望ましいネットワーク構造を決定することが必要となる。しかし、ベイズ推計法を用いてネットワーク構造を決定するためには極めて煩雑な計算が必要となる。このため、BNN法を用いてネットワーク構造を決定することは実用的でない。本研究では、NNモデルを用いて、土壤汚染浄化費用モデルの基本構造を決定する。その上で、BNNモデルを用いて、土壤汚染浄化費用リスクを評価するという手順を採用する。すなわち、BNNモデルは土壤汚染浄化費用リスクをベイズ予測するために用いることになる。本研究で提案する土壤汚染浄化費用リスクの評価手順は、図－2に示すように、1) 土壤汚染浄化費用予測を行うための核となるNNモデルを構築するプロセス（プロセス1）、2) NNモデルのネットワーク構造を所与として、BNN法により土壤汚染浄化費用リスクを評価するプロセス（プロセス2）で構成される。プロセス1では、過去の土壤汚染の浄化工事において、地籍調査、表土調査の結果と土壤汚染浄化費用の実績サンプルをデータベース化し、これらのデータを用いてNNモデルを作成する。作成したNNモデルを用いて、BNNモデルのネットワーク構造を確定する。プロセス2では、地籍・表土調査結果（以下、地点特性と呼ぶ）を用いて、当該地点における土壤汚染浄化費用リスクをBNNモデルにより評価する。その際、プロセス1で求めたNNモデルのネットワーク構造を所与として、BNN法によりNNモデルのパラメータの事後分布を推計するとともに、土壤汚染浄化費用リスクを評価することとなる。当然のことながら、当該地点における土壤汚染浄化費用の実績値が得られれば、この実測サンプルをデータベースに追加し、BNNモデル（もしくは、NNモデル）の更新を行うことになる。

3 BNNモデルの定式化

3.1 ベイズ予測問題

ベイズ統計学の詳細に関しては参考文献^{12),13)}に譲るが，ここでは読者の便宜を図るために，ベイズ推計法とベイズ予測問題を簡単に説明しておく．ベイズ推計法では，パラメータの事前分布と，実測サンプルを用いて定義される尤度関数を用いて，パラメータの事後分布を推計することになる．いま， M 個の実測サンプル $(\xi^1, \dots, \xi^M)$ が得られたとしよう．第 k ($k = 1, \dots, M$)番目の実測サンプルを $\xi^k = (y^k, \mathbf{x}^k)$ と表そう．ここに， y^k は実測サンプル k の土壤汚染浄化費用であり，実際の土壤汚染浄化工事のために支出された金額で表示される．一方， $\mathbf{x}^k = (x_1^k, \dots, x_n^k)$ は n 個の変数 x_i^k ($i = 1, \dots, n$)で記述される実測サンプル k の地点特性であり，土壤汚染浄化工事に先立って実施された地籍調査，表土調査の結果を表している．ここで，尤度関数を $\mathcal{L}(\xi^1, \dots, \xi^M | \theta)$ と表そう． θ は未知パラメータベクトルである．ここで， θ が確率変数で，事前確率密度関数 $\pi(\theta)$ に従うと仮定しよう．この時，実測サンプル $\xi^1, \dots, \xi^M$ を得たときに未知パラメータ θ の事後確率密度関数 $\pi(\theta | \xi^1, \dots, \xi^M)$ はベイズの定理より，

$$\pi(\theta | \xi^1, \dots, \xi^M) = \frac{\mathcal{L}(\xi^1, \dots, \xi^M | \theta) \pi(\theta)}{\int_{\Theta} \mathcal{L}(\xi^1, \dots, \xi^M | \theta) \pi(\theta) d\theta} \quad (3.1)$$

と表すことができる．ただし， Θ はパラメータ空間である．この時， $\pi(\theta | \xi^1, \dots, \xi^M)$ は

$$\pi(\theta | \xi^1, \dots, \xi^M) \propto \mathcal{L}(\xi^1, \dots, \xi^M | \theta) \pi(\theta) \quad (3.2)$$

と表すことができる． $\propto$ は比例を意味する．ここに，式(3.1)の分母

$$u(\xi) = \int_{\Theta} \mathcal{L}(\xi^1, \dots, \xi^M | \theta) \pi(\theta) d\theta \quad (3.3)$$

は事後分布 $\pi(\theta | \xi^1, \dots, \xi^M)$ の基準化定数，あるいは事前予測分布と呼ぶ．ベイズ推計は，1) 事前の経験情報などに基づき，事前確率密度関数 $\pi(\theta)$ を設定する．2) 過去に獲得した実測サンプル $(\xi^1, \dots, \xi^M)$ に基づいて尤度関数 $\mathcal{L}(\xi^1, \dots, \xi^M | \theta)$ を定義する．さらに，3) ベイズの定理(3.1)に基づいて事前確率密度関数を修正し，パラメータ θ に関する事後確率密度関数 $\pi(\theta | \xi^1, \dots, \xi^M)$ を求めるという手順をとる．最尤推定法と異なり，未知パラメータ θ の確率分布が，事後分布として求まる点にベイズ推計法の特徴がある．

土壤汚染浄化費用リスクのベイズ予測問題では、未知パラメータ θ を推計するだけでなく、新しい実測サンプル $\mathbf{x}^{M+1}$ を用いて、土壤汚染浄化費用 y^{M+1} の確率分布を予測することが問題となる。ベイズ予測問題は、過去の実測サンプル $(\xi^1, \dots, \xi^M)$ と、土壤汚染浄化費用を予測する地点のインプット情報 $\mathbf{x}^{M+1}$ に基づいて、実測サンプル $M+1$ の土壤汚染浄化費用 y^{M+1} の条件付き確率分布 $p(y^{M+1} | \mathbf{x}^{M+1}, \xi^1, \dots, \xi^M)$ を求める問題として記述できる。いま、地点特性一タベクトル $\mathbf{x}^{M+1}$ を与件とする土壤汚染浄化費用の条件付き確率密度関数を $p(y^{M+1} | \mathbf{x}^{M+1}, \theta)$ と表そう。以下、この条件付き確率密度関数を土壤汚染浄化費用モデルと呼ぶ。この時、ベイズ予測問題は

$$\begin{aligned} p(y^{M+1} | \mathbf{x}^{M+1}, \xi^1, \dots, \xi^M) \\ = \int p(y^{M+1} | \mathbf{x}^{M+1}, \theta) \pi(\theta | \xi^1, \dots, \xi^M) d\theta \end{aligned} \quad (3.4)$$

と表すことができる。図－1に示した土壤汚染浄化費用リスクは、式(3.4)を用いて求めた $p(y^{M+1} | \mathbf{x}^{M+1}, \xi^1, \dots, \xi^M)$ を図示したものである。

3.2 NNモデル（プロセス1）

プロセス1では、土壤汚染浄化費用モデル $p(y^{M+1} | \mathbf{x}^{M+1}, \theta)$ をNNモデルとして定式化する。NNモデルは、多くの比較的単純で独立したユニットをもち、相互に高い連携と、同時並列的な計算が行えるようなモデル構造を有している。NNモデルは、インプット（入力）、中間、アウトプット（出力）の各層に組織化されている。フィードフォワード・ネットワークはインプットからアウトプットまでを一方向のみのシグナル（信号）の流れ（インプット層から中間層、アウトプット層に至る）でマップ（地図化）する。中間層、アウトプット層のおのこのユニットは受けたシグナルを伝達する伝達関数をもつ。インプット層は伝達関数をもっていないが、インプット・シグナルをネットワーク全体に分配する機能をもっている。ユニットを結ぶおのこのコネクションは重み係数をもち、その重み係数は通過するシグナルに修正を加える。

本研究では、1つのアウトプット・ユニットと $m+1$ 個の中間層と $n+1$ 個のインプット・ユニットをもつ3層のフィードフォワード・ネットワークを考える。インプット層はベクトル $\mathbf{x} = (x_0, x_1, \dots, x_n)$ で表される。ただし、 x_0 はバイアス項であり、すべての実測サンプルに対して、恒常的に $x_0 = 1$ が成立する。また、インプット $(x_1, \dots, x_n)$

は、各実測サンプルの地点特性と対応している．当面、実測サンプル番号を表す上付き添字 k の表記を省略する．中間層のアウトプットベクトルを $\mathbf{z} = (z_0, z_1, z_2, \dots, z_m)$ で表す． z_0 はバイアス項であり、恒常的に $z_0 = 1$ が成立する．さらに、最終アウトプットを y で表そう．

中間層ユニット j ($j = 1, \dots, m$)は、すべてのインプットとバイアス項の加重和を受け取り、アウトプットシグナルの z_j を生成する．すなわち、

$$\begin{aligned} z_j &= F(\mathbf{x}\boldsymbol{\beta}'_j) \\ &= F\left(\sum_{i=0}^n \beta_{ij}x_i\right), \quad (j = 1, 2, \dots, m) \end{aligned} \quad (3.5)$$

が成立する．ここで、 F は伝達関数であり、 x_i は i 番目のインプットシグナル、 β_{ij} は i 番目のインプットシグナルから j 番目の中間層ユニットへのコネクションの重み係数である．また、ベクトル $\boldsymbol{\beta}_j = (\beta_{1j}, \dots, \beta_{nj})$ は、中間ユニット j のネットワークの重み係数ベクトルである．もちろん、中間ユニット j はすべてのインプット変数 x_i ($i = 1, \dots, n$)を含む必要はない．中間ユニット j がインプット変数 x_i の情報を伝達しない場合は、 $\beta_{ij} = 0$ が成立していると考えればいい．さらに、アウトプット・ユニットは各中間層ユニットからのアウトプットシグナル z_j を受け取り、土壌汚染浄化費用の推定値 y をシグナルとして生成するようなシステム

$$y = G\left(\sum_{j=0}^m \alpha_j z_j\right) \quad (3.6)$$

としてモデル化できる．ここに、 α_j は j 番目の中間層ユニットのアウトプットに対する重み係数である．式(3.5)を式(3.6)に代入すれば、NNモデルを

$$\begin{aligned} y &= f(\mathbf{x}, \boldsymbol{\beta}) \\ &= G\left(\alpha_0 + \sum_{j=1}^m \alpha_j F\left(\sum_{i=0}^n \beta_{ij}x_i\right)\right) \end{aligned} \quad (3.7)$$

と表現できる．ただし、 $\mathbf{x} = (x_0, \dots, x_n)$ はインプットベクトルである．また、 $\boldsymbol{\beta}$ はNNモデル全体の重み係数ベクトルであり、 $\boldsymbol{\beta} = (\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_m, \beta_{01}, \beta_{11}, \dots, \beta_{n1}, \beta_{02}, \dots, \beta_{n2}, \dots, \beta_{0m}, \beta_{1m}, \dots, \beta_{nm})$ で定義される．ここで、記述の便宜を図るために、重み係数ベクトルを $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p, \dots, \beta_s)$ と書き改めよう．ここに、 β_1 は α_0 に、 β_s は β_{nm} に対応している．また、 G, F はいくつかの関数形を採用することができるが、のちの適用事例では次に示すロジスティック関数を採用する．

$$G(x) = F(x) = \frac{1}{1 + \exp(-x)} \quad (3.8)$$

3.3 BNNモデル（プロセス2）

3.(1)で言及したように、土壌汚染浄化費用モデルは、地点特性ベクトル $\mathbf{x}$ を用いて、土壌汚染評価費用の条件付き確率密度関数を表現するモデル $p(y | \mathbf{x}, \boldsymbol{\theta})$ として定義できる。いま、NNモデルを用いて土壌汚染評価費用モデルを定式化しよう。そこで、土壌汚染浄化費用 y がNNモデルのアウトプット $f(\mathbf{x}, \boldsymbol{\beta})$ を平均値とする対数正規密度関数

$$\begin{aligned} p(y | \mathbf{x}, \boldsymbol{\theta}) &= (2\pi\sigma^2 y^2)^{-\frac{1}{2}} \exp \left\{ -\frac{(\ln y - \ln f(\mathbf{x}, \boldsymbol{\beta}))^2}{2\sigma^2} \right\} \end{aligned} \quad (3.9)$$

に従うと仮定しよう。ただし、 $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$ であり、 σ^2 は対数正規分布の分散（未知パラメータ）である。また、土壌汚染浄化費用 y は区間 $[0, \infty)$ で定義される。さらに、 M 個の実測サンプル $\hat{\boldsymbol{\xi}}^k = (\hat{y}^k, \hat{\mathbf{x}}^k)$, $(k = 1, \dots, M)$ が得られたとしよう。記号「 $\hat{\cdot}$ 」は、実測値であることを表す。この時、実測サンプル $(\hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M)$ に対して、尤度関数を

$$\begin{aligned} \mathcal{L}(\hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M | \boldsymbol{\theta}) &= \prod_{k=1}^M p(\hat{y}^k | \hat{\mathbf{x}}^k, \boldsymbol{\theta}) \\ &= (2\pi\sigma^2)^{-\frac{M}{2}} \prod_{k=1}^M \frac{1}{\hat{y}^k} \\ &\quad \exp \left\{ \sum_{k=1}^M -\frac{(\ln \hat{y}^k - \ln f(\hat{\mathbf{x}}^k, \boldsymbol{\beta}))^2}{2\sigma^2} \right\} \end{aligned} \quad (3.10)$$

と定義できる。

つぎに、パラメータ $\boldsymbol{\theta}$ の事後分布を求めよう。そのために、まず重み係数ベクトル $\boldsymbol{\beta}$ 、および分散 σ^2 の事前分布を定義しよう。本研究では、BNNモデルの重み係数 β_p ($p = 1, \dots, s$)が、それぞれ独立に期待値 $\tilde{\beta}_p$ 、分散 $\tilde{\sigma}_p^2$ ($p = 1, \dots, s$)の正規確率密度関数 $\mathcal{N}(\tilde{\beta}_p, \tilde{\sigma}_p^2)$ に従うと仮定する。5.で示す適用事例では、事前分布の平均値 $\tilde{\beta}_p$ として、プロセス1で求めたNNモデルの重み係数 β_p の最尤推定値を用いている。厳密には、重み係数 β_p ($p = 1, \dots, s$)の分布の間に相関関係が存在するため、重み係数ベクトル $\boldsymbol{\beta}$ の事前分布を多次元正規密度関数で定式化することが望ましい。しかし、分散・共分散行列に関する先験的情報が存在しないため、本研究では重み係数 β_p ($p = 1, \dots, s$)間の独立性を仮定する。実測サンプル数が増加すれば、事前分布の影響が低減する。実測サンプルが蓄積されるにつれて、事前分布の独立性の仮定は、実用上問題ではなくなる。また、標準的なベイズ統計学では、ベイズ推

計の容易さを考慮して、事前分布と事後分布が同一の関数形で表されるような共役分布を事前分布として採用することが多い．しかし、BNNモデルでは、事前分布と事後分布の関数形が等しくなるような共役事前分布を見いだすことは不可能である．したがって、事前分布として、操作性を考えて技術者が扱いやすい確率分布を想定することとした．このように比較的単純な事前確率密度関数を想定することにより、プロセス 1 で求めた NN モデルに関する情報を、土壌汚染浄化費用リスク評価のための事前分布として表現することが可能となる．いま、重み係数 β_p ($p = 1, \dots, s$) の事前分布を、正規確率密度関数

$$\pi(\beta_p) = (2\pi\tilde{\sigma}_p^2)^{-\frac{1}{2}} \exp \left\{ -\frac{(\beta_p - \tilde{\beta}_p)^2}{2\tilde{\sigma}_p^2} \right\} \quad (3.11)$$

で表そう．MCMC法を用いる場合、事前分布の分散が小さい場合、サンプリング効率が低下することが知られている．ここでは、標準偏差が $\tilde{\beta}_p$ に比例すると考え、

$$\tilde{\sigma}_p = \alpha \cdot \tilde{\beta}_p \quad (3.12)$$

と設定した．これにより、重み係数 β_p の事前分布の分散が極度に小さくなることが防止できる．当然のことながら、実測サンプル数が増加するに従って、事前分布の分散が土壌汚染浄化費用リスクに及ぼす影響は次第に低下していく．以上の結果、重み係数 β_p ($p = 1, \dots, s$) の事前分布は

$$\pi(\beta_p) = (2\pi\alpha^2\tilde{\beta}_p^2)^{-\frac{1}{2}} \exp \left\{ -\frac{(\beta_p - \tilde{\beta}_p)^2}{2\alpha^2\tilde{\beta}_p^2} \right\} \quad (3.13)$$

と表される．一方、分散 σ^2 は土壌汚染浄化費用リスクを評価するための重要なパラメータである． σ^2 に関しては、可能な限り実測サンプルにおける土壌汚染浄化費用リスクに関する情報を利用することとし、先験的情報を用いないこととした．そこで、 σ^2 の事前分布として、Jeffreys の無情報事前分布¹⁷⁾

$$\pi(\sigma^2) \propto \sigma^{-2} \quad (3.14)$$

を採用した．なお、Jeffreys 型無情報事前分布は定義区間上で積分しても 1 にならないので非正則な事前分布と呼ばれる．非正則な事前分布は確率密度関数ではないが、非正則な事前分布を用いてベイズ推計ルールにより事後分布を求めることができる^{16),17)}．以上で定義した尤度関数と事前分布を用いれば、重み係数ベクト

ル $\boldsymbol{\beta}$ の事後周辺分布は

$$\begin{aligned}
& \pi(\boldsymbol{\beta} \mid \sigma^2, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) \\
& \propto \mathcal{L}(\hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M \mid \boldsymbol{\beta}, \sigma^2) \cdot \prod_{p=1}^s \pi(\beta_p) \\
& \propto \exp \left\{ \sum_{p=1}^s -\frac{(\beta_p - \tilde{\beta}_p)^2}{2\alpha^2 \tilde{\beta}_p^2} \right. \\
& \quad \left. - \sum_{k=1}^M \frac{(\ln \hat{y}^k - \ln f(\hat{\boldsymbol{x}}^k, \boldsymbol{\beta}))^2}{2\sigma^2} \right\}
\end{aligned} \tag{3.15}$$

と表される．一方，推定誤差 σ^2 の事後周辺分布は

$$\begin{aligned}
& \pi(\sigma^2 \mid \boldsymbol{\beta}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) \\
& \propto \mathcal{L}(\hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M \mid \boldsymbol{\beta}, \sigma^2) \cdot \pi(\sigma^2) \\
& \propto (\sigma^2)^{-(\frac{M}{2}+1)} \exp \left\{ - \sum_{k=1}^M \frac{(\ln \hat{y}^k - \ln f(\hat{\boldsymbol{x}}^k, \boldsymbol{\beta}))^2}{2\sigma^2} \right\}
\end{aligned} \tag{3.16}$$

と表すことができる．

4 土壌汚染浄化費用リスクの評価

4.1 リスク評価の方法

ベイズ統計学を用いて事後分布を求める際、式(3.3)に示す基準化定数に関する情報が必要となる。しかし、多くの場合、基準化定数を解析的に求めることは不可能であり、このことがベイズ統計学を実務へ適用する大きな障害になっていた。しかし、近年MCMC法¹⁶⁾がベイズ統計学の分野に導入され、多重数値積分により基準化定数を求めなくても、効率的に事後分布を求めることが可能となった。その結果、ベイズ推計法の適用範囲は大幅に拡大したと考えることができる。すでに、MCMC法を用いたベイズ推計法に関して、いくつかの研究が蓄積されている¹⁶⁾。代表的なMCMC法として、ギブスサンプリング (Gibbs sampling) 法、メトロポリス・ヘイスティングス (Metropolis-Hastings : MHと略す) 法等が提案されている¹⁶⁾。この内、ギブスサンプリングはもともと画像復元のアルゴリズム¹⁸⁾として知られていたが、ベイズ推計法における事後分布の推計に応用された¹⁹⁾。ギブスサンプリング法、MH法は、いずれも事後確率密度関数 $\pi(\theta|\hat{\xi}^1, \dots, \hat{\xi}^M)$ を直接求めることが難しい場合に、各パラメータの条件付き事後確率密度関数を用いて、反復的にパラメータ β のサンプルを乱数発生させることにより、事後分布からパラメータサンプルを獲得する方法である。しかし、ギブスサンプリングにより事後確率分布から効率的にパラメータサンプルを発生させるためには、対数事後確率密度関数が凹関数であることが必要となる²⁰⁾。しかし、BNNモデルの場合、対数事後確率密度関数が大域的に凹である保証がない。このため、本研究では対数事後確率密度関数の凹性が保証されなくても、パラメータの条件付き事後確率密度関数から容易にサンプルを発生させることができるMH法を採用することとした。

4.2 MHアルゴリズム

MHアルゴリズムは、事後分布 $\pi(\theta|\hat{\xi}^1, \dots, \hat{\xi}^M)$ から、直接パラメータサンプルを発生させることが難しい場合に、これを近似するような代替的な分布からサンプリングを行い、それに基づいて本来の事後分布からのサンプルを求める方法である。近似分布として、サンプルを乱数発生させやすい確率分布が採用される。事後分布は本来サンプリングを行うべき分布であり目標分布と呼ばれる。近似分布から乱数を発生させ、本節で述べるような修正を行うステップを経ることにより、

最終的に目標分布からサンプルを発生させることが可能となる．近似分布は目標分布から抽出するサンプルの候補を提案・生成することから提案分布と呼ばれる．

各パラメータの初期値 $\boldsymbol{\beta}^{(0)} = (\beta_1^{(0)}, \dots, \beta_s^{(0)})$, $\sigma^{2(0)}$ を設定する．さらに，以下の手順で順次パラメータベクトルのサンプル $(\boldsymbol{\beta}^{(i)}, \sigma^{2(i)})$, $(i = 1, 2, \dots)$ を抽出する．いま， $(\boldsymbol{\beta}^{(i-1)}, \sigma^{2(i-1)})$ が与えられたとしよう．この時，提案分布 q_p ($p = 1, \dots, s$), t を用いてサンプル候補 $\bar{\boldsymbol{\beta}}, \bar{\sigma}^2$ を発生させる問題を考えよう．まず，パラメータサンプル $\beta_p^{(i)}$ を与件とした条件付分布 $q_p(\beta_p^{(i)})$ （提案分布）を用いて新しいサンプルパラメータ $\bar{\beta}_p$ をランダムサンプリングしよう．いま， $\bar{\beta}_p$ ($p = 1, \dots, s$)に関する提案分布を $\beta_p^{(i-1)}$ を平均， $\alpha^2 \beta_p^{(i-1)^2}$ を分散とする正規分布

$$\begin{aligned} \bar{\beta}_p &\sim q_p(\beta_p^{(i-1)}) \\ &= \frac{1}{\sqrt{2\pi\alpha\beta_p^{(i-1)}}} \exp \left\{ -\frac{(\bar{\beta}_p - \beta_p^{(i-1)})^2}{2\alpha^2 \beta_p^{(i-1)^2}} \right\} \end{aligned} \quad (4.1)$$

により表現する．提案された $\bar{\beta}_p$ は目標分布からのサンプルではないので，目標分布からのズレを修正するために，確率

$$\begin{aligned} \mathcal{A}(\boldsymbol{\beta}^{(i-1)}, \bar{\boldsymbol{\beta}} \mid \sigma^{2(i-1)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) \\ = \min \left[\frac{\pi(\bar{\boldsymbol{\beta}} \mid \sigma^{2(i-1)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M)}{\pi(\boldsymbol{\beta}^{(i-1)} \mid \sigma^{2(i-1)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M)}, 1 \right] \end{aligned} \quad (4.2)$$

で受容し， $\boldsymbol{\beta}^{(i)} = \bar{\boldsymbol{\beta}}$ とする．もし，棄却された場合には $\boldsymbol{\beta}^{(i)} = \boldsymbol{\beta}^{(i-1)}$ とし，もとの点にとどまる．実際のプログラミングにおいては区間 $(0, 1)$ 上の一様乱数 $u \sim U(0, 1)$ を発生させ， $u \leq \mathcal{A}$ のとき $\bar{\boldsymbol{\beta}}$ を受容し，それ以外るとき棄却すればよい．つぎに， $\bar{\sigma}^2$ の新しい候補点を平均を $\sigma^{2(i-1)}$ とする指数分布

$$\begin{aligned} \bar{\sigma}^2 &\sim t(\sigma^{2(i-1)}) \\ &= \frac{1}{\sigma^{2(i-1)}} \exp \left\{ -\frac{\bar{\sigma}^2}{\sigma^{2(i-1)}} \right\} \end{aligned} \quad (4.3)$$

で提案する．提案分布 t は，それぞれ $\sigma^{2(i-1)}$ を与件とした $\bar{\sigma}^2$ の条件付確率密度関数である．目標分布からの乖離を補正するために，確率

$$\begin{aligned} \mathcal{B}(\sigma^{2(i-1)}, \bar{\sigma}^2 \mid \boldsymbol{\beta}^{(i)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) \\ = \min \left[\frac{\pi(\bar{\sigma}^2 \mid \boldsymbol{\beta}^{(i)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M)}{\pi(\sigma^{2(i-1)} \mid \boldsymbol{\beta}^{(i)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M)}, 1 \right] \end{aligned} \quad (4.4)$$

で受容し, $\sigma^{2(i)} = \bar{\sigma}^2$ とする. 棄却された場合には $\sigma^{2(i)} = \sigma^{2(i-1)}$ とする.

以上の手順をまとめると次のようになる.

- 1) **ステップ 1** 初期パラメータ値 $\boldsymbol{\beta}^{(0)} = (\beta_1^{(0)}, \dots, \beta_s^{(0)})$, $\sigma^{2(0)}$ を設定する. ただし, $\boldsymbol{\beta}^{(0)}, \sigma^{2(0)}$ として, NN モデルのパラメータの最尤推定値を用いる. $i = 1$ に設定し, ステップ 2 に進む.
- 2) **ステップ 2** パラメータサンプル $\boldsymbol{\beta}^{(i-1)} = (\beta_1^{(i-1)}, \dots, \beta_s^{(i-1)})$ を与件とした提案分布(式(4.1))を用いて新しいサンプルパラメータ $\bar{\boldsymbol{\beta}} = (\bar{\beta}_1, \dots, \bar{\beta}_s)$ を提案する.
- 3) **ステップ 3** 受容確率

$$\begin{aligned} & \mathcal{A}(\boldsymbol{\beta}^{(i-1)}, \bar{\boldsymbol{\beta}} \mid \sigma^{2(i-1)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) \\ &= \min \left[\exp \left(\sum_{p=1}^s \frac{(\beta_p^{(i-1)} - \bar{\beta}_p)^2 - (\bar{\beta}_p - \tilde{\beta}_p)^2}{2\alpha^2 \tilde{\beta}_p^2} + \Phi \right), 1 \right] \end{aligned} \quad (4.5)$$

を求める. ただし,

$$\Phi = \sum_{k=1}^M \left\{ \frac{\left(\ln \frac{\hat{y}^k}{f(\hat{\mathbf{x}}^k, \boldsymbol{\beta}_p^{(i-1)})} \right)^2 - \left(\ln \frac{\hat{y}^k}{f(\hat{\mathbf{x}}^k, \bar{\boldsymbol{\beta}}_p)} \right)^2}{2\sigma^{2(i-1)}} \right\}$$

である.

- 4) **ステップ 4** 一様分布 $u \sim U(0, 1)$ を発生させ, $\boldsymbol{\beta}^{(i)}$ を次式で決定する.

$$\boldsymbol{\beta}^{(i)} = \begin{cases} \bar{\boldsymbol{\beta}} & u \leq \mathcal{A} \text{ の場合} \\ \boldsymbol{\beta}^{(i-1)} & \text{その他の場合} \end{cases} \quad (4.6)$$

- 5) **ステップ 5** 新しい分散候補 $\bar{\sigma}^2$ を平均 $\sigma^{2(i-1)}$ とする指数分布(式(4.3))に従ってランダムサンプルする.

- 6) **ステップ 6** 受容確率

$$\begin{aligned} & B(\sigma^{2(i-1)}, \bar{\sigma}^2 \mid \boldsymbol{\beta}^{(i)}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) \\ &= \min \left[\left(\frac{\bar{\sigma}^2}{\sigma^{2(i-1)}} \right)^{-(\frac{M}{2}+1)} \exp \Theta, 1 \right] \end{aligned} \quad (4.7)$$

を求める. ただし,

$$\Theta = \sum_{k=1}^M \left\{ \frac{\left(\ln \frac{\hat{y}^k}{f(\hat{\mathbf{x}}^k, \boldsymbol{\beta}^{(i)})} \right)^2}{2\sigma^{2(i-1)}} - \frac{\left(\ln \frac{\hat{y}^k}{f(\hat{\mathbf{x}}^k, \bar{\boldsymbol{\beta}}^{(i)})} \right)^2}{2\bar{\sigma}^2} \right\}$$

である.

7) **ステップ 7** 一様分布 $v \sim U(0, 1)$ を発生させ、 $\sigma^{2(i)}$ を次式で決定する．

$$\sigma^{2(i)} = \begin{cases} \bar{\sigma}_2 & v \leq \mathcal{B} \text{ の場合} \\ \sigma^{2(i-1)} & \text{その他の場合} \end{cases} \quad (4.8)$$

8) **ステップ 8** 十分大きな $\underline{N}$ に対して $(\beta^{(i)}, \sigma^{2(i)})$, $i = \underline{N}, \underline{N}+1, \dots$ を記録する． $i = i+1$, $p = 1$ としてステップ 2 に戻る． $i = \bar{N}$ ($\bar{N} > \underline{N}$) に到達したらアルゴリズムを終了する．ただし、 $\bar{N}$ は、あらかじめ設定したサンプリング回数である．

4.3 事後分布に関する統計量

MH法によって得られたサンプルに基づいて、パラメータ β に関する統計的性質を分析することができる．ベイズ統計学を用いた事後分布に関する統計量に関しては、津田¹⁴⁾が整理している．ここでは、読者の便宜を図るため、既存論文¹⁴⁾との重複はあるものの、事後分布に関する統計量を以下でとりまとめておく．MH法を用いた場合、パラメータの事後確率密度関数 $\pi(\theta | \hat{\xi}^1, \dots, \hat{\xi}^M)$ を解析的な関数として表現することはできない．得られたサンプルを用いてノンパラメトリックに分布関数や密度関数を推計することとなる．いま、MHサンプリングから得られたサンプルを $\theta^{(i)} = (\beta^{(i)}, \sigma^{2(i)})$ ($i = 1, \dots, \bar{N}$) と表そう．この内、最初の $\underline{N}$ 個のサンプルはMHサンプリングの収束過程からのサンプルと考え、サンプル集合から除去しよう．その上で、パラメータサンプルの添字集合を $\mathcal{M} = \{\underline{N}+1, \dots, \bar{N}\}$ と定義する．この時、パラメータ β_p ($p = 1, \dots, s$) に関する確率分布関数 $F_p(\beta_p)$ は次式で近似できる．

$$F_p(\beta_p) = \frac{\#\{\beta_p^{(i)} \leq \beta_p, i \in \mathcal{M}\}}{\bar{N} - \underline{N}} \quad (4.9)$$

ただし、 $\#\{\beta_p^{(i)} \leq \beta_p, i \in \mathcal{M}\}$ は論理式 $\{\beta_p^{(i)} \leq \beta_p, i \in \mathcal{M}\}$ が成立するサンプルの総数である．同様に、パラメータ σ^2 の確率分布関数 $G(\sigma^2)$ も

$$G(\sigma^2) = \frac{\#\{\sigma^{2(i)} \leq \sigma^2, i \in \mathcal{M}\}}{\bar{N} - \underline{N}} \quad (4.10)$$

と表すことができる．また、パラメータ β_p の事後分布の平均 $\tilde{\mu}(\beta_p)$ 、分散 $\tilde{v}(\beta_p)$ は標本平均、標本分散を用いて表される．すなわち、

$$\tilde{\mu}(\beta_p) = \sum_{i=\underline{N}+1}^{\bar{N}} \frac{\beta_p^{(i)}}{\bar{N} - \underline{N}} \quad (4.11a)$$

$$\tilde{v}(\beta_p) = \sum_{i=\underline{N}+1}^{\bar{N}} \frac{\{\beta_p^{(i)} - \tilde{\mu}(\beta_p)\}^2}{\bar{N} - \underline{N}} \quad (4.11b)$$

と表される．また，パラメータ σ^2 の平均 $\tilde{\mu}(\sigma^2)$ ，分散 $\tilde{v}(\sigma^2)$ も

$$\tilde{\mu}(\sigma^2) = \sum_{i=\underline{N}+1}^{\overline{N}} \frac{\sigma^{2(i)}}{\overline{N} - \underline{N}} \quad (4.12a)$$

$$\tilde{v}(\sigma^2) = \sum_{i=\underline{N}+1}^{\overline{N}} \frac{\{\sigma^{2(i)} - \tilde{\mu}(\sigma^2)\}^2}{\overline{N} - \underline{N}} \quad (4.12b)$$

と表される．パラメータ β_p, σ^2 の $100(1 - 2\alpha)\%$ 信頼区間に関しては，標本順序統計量

$$\begin{aligned} \underline{\beta}_p(\alpha) &= \arg \sup_{\beta_p^{(i^*)}} \\ &\quad \left\{ \frac{\#\{\beta_p^{(i)} \leq \beta_p^{(i^*)}, i \in \mathcal{M}\}}{\overline{N} - \underline{N}} < \alpha \right\} \end{aligned} \quad (4.13a)$$

$$\begin{aligned} \overline{\beta}_p(\alpha) &= \arg \inf_{\beta_p^{(i^{**})}} \\ &\quad \left\{ \frac{\#\{\beta_p^{(i)} \geq \beta_p^{(i^{**})}, i \in \mathcal{M}\}}{\overline{N} - \underline{N}} < \alpha \right\} \end{aligned} \quad (4.13b)$$

$$\begin{aligned} \underline{\sigma}^2(\alpha) &= \arg \sup_{\sigma^{2(i^*)}} \\ &\quad \left\{ \frac{\#\{\sigma^{2(i)} \leq \sigma^{2(i^*)}, i \in \mathcal{M}\}}{\overline{N} - \underline{N}} < \alpha \right\} \end{aligned} \quad (4.13c)$$

$$\begin{aligned} \overline{\sigma}^2(\alpha) &= \arg \inf_{\sigma^{2(i^{**})}} \\ &\quad \left\{ \frac{\#\{\sigma^{2(i)} \geq \sigma^{2(i^{**})}, i \in \mathcal{M}\}}{\overline{N} - \underline{N}} < \alpha \right\} \end{aligned} \quad (4.13d)$$

を用いて $\underline{\beta}_p(\alpha) \leq \beta_p \leq \overline{\beta}_p(\alpha)$ ($p = 1, \dots, s$), $\underline{\sigma}^2(\alpha) \leq \sigma^2 \leq \overline{\sigma}^2(\alpha)$ と定義できる．なお，「arg」は，操作「sup」，あるいは「inf」を達成する変数値を指示する記号である．式(4.13b)の場合，右辺が満足するような $\beta_p^{(i^*)}$ の上界値を指示している．

MH法では，初期パラメータ $\theta^{(0)}$ が不変分布である事後分布からのサンプルである保証はない．いま，MHサンプリングを合計 $\overline{N}$ 回繰り返した場合を考えよう． $\overline{N}$ 個のサンプルの内，最初の $\underline{N}$ 個のサンプル $\theta^{(i)}$ ($i = 1, \dots, \underline{N}$)を事後分布に収束する過程からのサンプリングと考える．その上で，第 $\underline{N}+1$ 回以降のサンプルをとりあげよう． $\underline{N}+1$ 以降のサンプルが，不変分布である事後分布からのサンプルであるかどうかをGewekeの方法²³⁾により仮説検定できる．いま，パラメータのMHサンプル $\theta^{(i)}$ ($i = \underline{N}+1, \dots, \overline{N}$)の中から，最初の N_1 個と最後の N_2 個のサンプルをとりあげよう．この時，パラメータ β_p ($p = 1, \dots, s$)の不変分布への収束を判断するためのGeweke

検定統計量は,

$$\begin{aligned}
Z_{\beta_p} &= \frac{\bar{\beta}_p^1 - \bar{\beta}_p^2}{\sqrt{\nu_1(\beta_p)^2 + \nu_2(\beta_p)^2}} \sim \mathcal{N}(0, 1) \\
\bar{\beta}_p^1 &= \frac{\sum_{i=N+1}^{N+N_1} \beta_p^{(i)}}{N_1} \quad \bar{\beta}_p^2 = \frac{\sum_{i=\bar{N}-N_2+1}^{\bar{N}} \beta_p^{(i)}}{N_2} \\
\nu_1^2(\beta_p) &= \frac{2\pi \hat{f}_p^1(0)}{N_1} \quad \nu_2^2(\beta_p) = \frac{2\pi \hat{f}_p^2(0)}{N_2}
\end{aligned} \tag{4.14}$$

と定義できる. ただし, $f_p^i(x)$ ($i = 1, 2$) はスペクトル密度関数であり, $2\pi f_p^i(0)$ の推定値は

$$\begin{aligned}
2\pi \hat{f}_p^i(0) &= \hat{\omega}_0^i + 2 \sum_{s=1}^q w(s, q) \hat{\omega}_j^i \\
\hat{\omega}_j^1 &= N_1^{-1} \sum_{g=N+s+1}^{N+N_1} (\beta_p^{(g)} - \bar{\beta}_p^1)(\beta_p^{(g-s)} - \bar{\beta}_p^1) \\
\hat{\omega}_j^2 &= N_2^{-1} \sum_{g=\bar{N}-N_2+s+1}^{\bar{N}} (\beta_p^{(g)} - \bar{\beta}_p^2)(\beta_p^{(g-s)} - \bar{\beta}_p^2) \\
w(s, q) &= 1 - \frac{s}{q+1}
\end{aligned} \tag{4.15}$$

として求まる^{21), 24)}. 同様に, パラメータ σ^2 についても Geweke 検定統計量を

$$\begin{aligned}
Z_{\sigma^2} &= \frac{(\bar{\sigma}^2)_1 - (\bar{\sigma}^2)_2}{\sqrt{\nu_1(\sigma^2)^2 + \nu_2(\sigma^2)^2}} \sim \mathcal{N}(0, 1) \\
(\bar{\sigma}^2)_1 &= \frac{\sum_{i=N+1}^{N+N_1} \sigma^{2(i)}}{N_1} \quad (\bar{\sigma}^2)_2 = \frac{\sum_{i=\bar{N}-N_2+1}^{\bar{N}} \sigma^{2(i)}}{N_2} \\
\nu_1(\sigma^2)^2 &= \frac{2\pi \hat{f}_{\sigma^2}^1(0)}{N_1} \quad \nu_2(\sigma^2)^2 = \frac{2\pi \hat{f}_{\sigma^2}^2(0)}{N_2}
\end{aligned} \tag{4.16}$$

と定義する. ここで, パラメータ β_p の不変分布への収束性に関する帰無仮説 $H_0(\beta_p)$ と対立仮説 $H_1(\beta_p)$ を

$$\begin{cases} H_0(\beta_p) : |Z_{\beta_p}| \leq z_{\alpha/2} \\ H_1(\beta_p) : |Z_{\beta_p}| > z_{\alpha/2} \end{cases} \tag{4.17}$$

と設定しよう. ただし, $z_{\alpha/2}$ は帰無仮説を棄却するための臨界的な値である. 有意水準 $\alpha\%$ で帰無仮説を仮説検定する場合, $z_{\alpha/2}$ は $\alpha/2\% = 1 - \Phi(z_{\alpha/2})$ を満足する値として定義できる. ただし, $\Phi(z)$ は標準正規分布の分布関数である. 同様に, σ^2 についても, 仮説検定を行うことができる. 最後に, BNN モデルの推計精度はベイズファ

クター

$$B = \frac{\mathcal{S}(\hat{\xi}^1, \dots, \hat{\xi}^M)}{\mathcal{S}_\phi(\hat{\xi}^1, \dots, \hat{\xi}^M)} \quad (4.18)$$

で与えられる．ここに， $\mathcal{S}(\hat{\xi}^1, \dots, \hat{\xi}^M)$ はBNNモデルを用いた場合の周辺尤度， $\mathcal{S}_\phi(\hat{\xi}^1, \dots, \hat{\xi}^M)$ はnull情報の下での周辺尤度を表す．周辺尤度 $\mathcal{S}(\hat{\xi}^1, \dots, \hat{\xi}^M)$ は

$$\mathcal{S}(\hat{\xi}^1, \dots, \hat{\xi}^M) = \int \tilde{\mathcal{L}}(\hat{\xi}^1, \dots, \hat{\xi}^M | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (4.19a)$$

$$\mathcal{S}_\phi(\hat{\xi}^1, \dots, \hat{\xi}^M) = \tilde{\mathcal{L}}(\hat{\xi}^1, \dots, \hat{\xi}^M | \mathbf{o}) \quad (4.19b)$$

と定義できる．ただし， $\tilde{\mathcal{L}}(\hat{\xi}^1, \dots, \hat{\xi}^M | \boldsymbol{\theta})$ はパラメータベクトル $\boldsymbol{\theta}$ の下で実測サンプル $\hat{\xi}^1, \dots, \hat{\xi}^M$ に関する同時生起確率密度関数を表し，尤度関数 $\tilde{\mathcal{L}}(\boldsymbol{\theta} | \hat{\xi}^1, \dots, \hat{\xi}^M)$ を $(\hat{\xi}^1, \dots, \hat{\xi}^M)$ に関する確率密度関数と解釈することにより得られる．

4.4 土壌汚染浄化費用リスクの評価

新しい調査地点 $M+1$ の地籍調査，表土調査を実施して，地点特性の実測値 $\hat{\mathbf{x}}^{M+1}$ が得られたとしよう．その上で，調査地点 $M+1$ の土壌汚染浄化費用 y^{M+1} のリスクを評価する問題を考える．いま，MH法により，パラメータサンプル $\boldsymbol{\theta}^{(i)} = (\boldsymbol{\beta}^{(i)}, \sigma^{2(i)})$ ， $(i = \underline{N} + 1, \dots, \overline{N})$ が得られたとしよう．これらのパラメータサンプルにより，パラメータ $\boldsymbol{\theta}$ の事後分布 $\pi(\boldsymbol{\theta} | \hat{\xi}^1, \dots, \hat{\xi}^M)$ を離散近似できる．さらに，パラメータサンプル $\boldsymbol{\theta}^{(i)} = (\boldsymbol{\beta}^{(i)}, \sigma^{2(i)})$ を所与とした時，調査地点 $M+1$ の土壌汚染浄化費用 y^{M+1} はNNモデルのアウトプット $f(\hat{\mathbf{x}}^{M+1}, \boldsymbol{\beta}^{(i)})$ を平均値とする対数正規密度関数

$$p(y^{M+1} | \hat{\mathbf{x}}^{M+1}, \boldsymbol{\theta}^{(i)}) = \{2\pi\sigma^{2(i)}(y^{M+1})^2\}^{-\frac{1}{2}} \cdot \exp\left\{-\frac{(\ln y^{M+1} - \ln f(\hat{\mathbf{x}}^{M+1}, \boldsymbol{\beta}^{(i)}))^2}{2\sigma^{2(i)}}\right\} \quad (4.20)$$

に従う．この時，式(3.4)より，調査地点 $M+1$ の土壌汚染浄化費用リスクは，条件付き確率密度関数

$$\begin{aligned} p(y^{M+1} | \hat{\mathbf{x}}^{M+1}, \hat{\xi}^1, \dots, \hat{\xi}^M) \\ = \int p(y^{M+1} | \hat{\mathbf{x}}^{M+1}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | \hat{\xi}^1, \dots, \hat{\xi}^M) d\boldsymbol{\theta} \end{aligned} \quad (4.21)$$

を用いて表現できる．MH法で求めたパラメータサンプルを用いて，土壌汚染浄化リスク $p(y^{M+1} | \hat{\mathbf{x}}^{M+1}, \hat{\xi}^1, \dots, \hat{\xi}^M)$ を，以下の手順で離散近似することができる．す

なわち，パラータサンプル $\boldsymbol{\theta}^{(i)}$ を所与とした確率密度関数(4.20)から， ρ ($\rho \geq 1$)個の費用サンプル $\mathbf{y}^{(i)} = (y_1^{(i)}, \dots, y_\rho^{(i)})$ を抽出する．ついで，サンプルパラメータ集合 $\{\boldsymbol{\theta}^{(i)} \ (i \in \mathcal{M})\}$ に対して，費用サンプル集合 $\{\mathbf{y}^{(i)} \ (i \in \mathcal{M})\}$ を定義する．最後に，土壤汚染処理費用を H 個の離散的カテゴリ $[y_{h-1}, y_h), (h = 1, \dots, H)$ に分割する．ただし， y_H は十分大きい正数であり， $0 = y_0 < y_1 < \dots < y_h < y_{h+1} < \dots < y_H < \infty$ が成立する．その上で，相対頻度 R_h ($h = 1, \dots, H$)を

$$R_h = \frac{\#\{y_{h-1} \leq y_t^{(i)} < y_h; y_t^{(i)} \in \Omega\}}{\#\{y_t^{(i)} \in \Omega\}} \quad (4.22)$$

と定義する．ただし， $\Omega = \{y_t^{(i)} \ (t = 1, \dots, \rho; i \in \mathcal{M})\}$ である．以上で求めた相対頻度 R_h ($h = 1, \dots, H$)は，土壤汚染費用リスクを表す確率密度関数 $p(y^{M+1} \mid \hat{\mathbf{x}}^{M+1}, \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M)$ の離散近似を与える．さらに，リスク管理指標としてVaR指標を用いよう．危険水準を $100 \cdot \alpha\%$ とするVaR指標は

$$\text{VaR}(\alpha) = \arg \inf_{y^*} \left\{ \frac{\#\{y_t^{(i)} \geq y^*; y_t^{(i)} \in \Omega\}}{\#\{y_t^{(i)} \in \Omega\}} < \alpha \right\} \quad (4.23)$$

と表すことができる．

4.5 BNNモデルのベイズ更新

ベイズ予測問題は過去の実測サンプル $\hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M$ のストックと，新しい調査地点に対して地籍調査，表土調査により得られた地点特性情報 $\hat{\mathbf{x}}^{M+1}$ に基づいて，その地点における土壤汚染浄化費用の確率分布を求める問題として定式化できる．その結果に基づいて，土地所有者が土地開発を決定し，実際に土地浄化工事を実施すれば，調査地点 $M+1$ の土壤汚染浄化費用の実績値 $\hat{y}^{M+1}$ を獲得できる．新しい実測サンプル $\hat{\boldsymbol{\xi}}^{M+1} = (\hat{y}^{M+1}, \hat{\mathbf{x}}^{M+1})$ を用いてBNNモデルを更新し，新しい調査地点 $M+2$ の土壤汚染浄化費用の予測精度向上に努めることが可能となる．このように新しい実測サンプルが付加されるごとに，BNNモデルの内容を更新することをBNNモデルのベイズ更新と呼ぶ．

ベイズ推計において，事前分布と事後分布が同一の関数形を持つような共役事前分布が存在する場合，新しく獲得した追加データを用いて，未知パラメータのベイズ推計値を容易に更新することができる．残念ながら，BNNモデルの場合，共役事

前分布を求めることが不可能である．このため，ベイズ更新を実施するためには，モデルの推計に用いる過去の実測サンプルをすべて蓄積しておかなければならない．いま，ベイズ更新の方法を説明するために， M 個の実測サンプル $(\hat{\xi}^1, \dots, \hat{\xi}^M)$ を用いてBNNモデルの重み係数ベクトルの事後周辺分布(3.15)と分散の事後周辺分布(3.16)が求まった場合を考えよう．その上で，第 $M+1$ 個目の実測サンプル $\hat{\xi}^{M+1}$ を用いて，未知パラメータの事後周辺分布を更新する問題を考えよう．未知パラメータの事後周辺分布を更新する場合，1) 事前分布を修正せずに，尤度関数 $\mathcal{L}(\theta | \hat{\xi}^1, \dots, \hat{\xi}^M)$ のみを修正する方法，2) 事前分布と尤度関数の双方を同時に修正する方法が考えられる．まず，尤度関数のみを修正する場合をとりあげよう．いま，新しい実測サンプルを加えた $M+1$ 個のデータベースが整備されたと考えよう．この時，重み係数ベクトル β の事後周辺分布は

$$\begin{aligned} \pi(\beta | \sigma^2, \hat{\xi}^1, \dots, \hat{\xi}^{M+1}) \\ \propto \exp \left\{ \sum_{p=1}^s -\frac{(\beta_p - \tilde{\beta}_p)^2}{2\alpha^2 \tilde{\beta}_p^2} - \sum_{k=1}^{M+1} \frac{(\ln \hat{y}^k - \ln f(\hat{\mathbf{x}}^k, \beta))^2}{2\sigma^2} \right\} \end{aligned} \quad (4.24)$$

と更新される．一方，推定誤差 σ^2 の事後周辺分布は

$$\begin{aligned} \pi(\sigma^2 | \beta, \hat{\xi}^1, \dots, \hat{\xi}^{M+1}) \\ \propto (\sigma^2)^{-(\frac{M}{2}+1)} \exp \left\{ - \sum_{k=1}^{M+1} \frac{(\ln \hat{y}^k - \ln f(\hat{\mathbf{x}}^k, \beta))^2}{2\sigma^2} \right\} \end{aligned} \quad (4.25)$$

と表すことができる．新しい調査地点 $M+2$ の土壤汚染浄化費用リスクを評価する場合には，更新された事後周辺分布(4.24),(4.25)を用いて，MH法により土壤汚染浄化費用の確率分布を求めることになる．さらに，事前分布も更新する場合を考えよう．この場合，重み係数ベクトル β の事前分布(3.13)の平均値ベクトル $\tilde{\beta}$ を修正することにより対応できる．本研究では最尤推定法により推定した重み係数ベクトルを，事前分布の平均値ベクトルとして採用している．このことは，事前分布を再定義するためにNNネットワークを再構成することを意味している．したがって，尤度関数のみを修正する場合と比較して，更新のための計算量は飛躍的に増加する．のちに，5.(3)において，2つのベイズ更新の方法が，土壤汚染浄化費用リスクの推計結果に及ぼす影響について考察することとする．

5 適用事例

5.1 適用事例の概略

わが国で土壤汚染対策が実施されるようになって日も浅く，土壤汚染に関するデータの蓄積は不十分である．さらに，土壤汚染浄化費用と，地籍調査や表土調査による地点特性の双方が同時に利用可能であるデータは極端に限られている．本研究では，過去に実施された土壤汚染浄化対策に関わる実績サンプルを用いる．研究対象として用いるデータベースには，掘削除去工法と原位置浄化法を採用した実測サンプルの双方が含まれている．2つの浄化工法を比較した場合，明らかに費用構造が異なるため，本研究では掘削除去工法を採用した実測サンプルのみを対象とする．本データベースでは，土壤汚染浄化工事に先立って実施された地籍調査，表土調査の結果も利用可能であり，表－1に示すような地点特性に関する情報が入手可能である．

BNNモデルの推計に用いることができる実測サンプル数は34個である．BNNモデルの推計に用いるデータとしては，必ずしも十分とは言えない．現時点におけるデータの蓄積状況を考慮すれば，本データベースを使用せざるを得ないのが実情である．そのため，BNNモデルを用いてベイズ予測する土壤汚染浄化費用リスクは，相当程度大きいことが予想される．今後，実測サンプルが蓄積されれば，土壤汚染浄化費用の推定リスクを縮減することが可能であろう．言い換えれば，現時点のデータ蓄積量を前提とすれば，簡易調査結果に基づいて土壤汚染浄化費用を見積もる時には，本適用事例で示す程度の土壤汚染浄化費用リスクが存在することを想定せざるを得ない．

5.2 BNNモデルの推計

図－2に示した方法に従って，まず通常のNNモデルの作成を試みた（プロセス1）．NNモデルの重み係数ベクトル β と分散 σ^2 を最尤法により学習した．式(3.9)より，対数尤度関数は

$$\begin{aligned}\ln \mathcal{L} &= \ln \prod_{k=1}^M p(y^k | \hat{\mathbf{x}}^k, \theta) \\ &= \iota - M \ln \sigma - \sum_{k=1}^n \frac{\{\ln \hat{y}^k - \ln f(\hat{\mathbf{x}}^k, \beta)\}^2}{2\sigma^2}\end{aligned}\quad (5.1)$$

と定式化できる．ただし， ι は定数である．対数尤度関数を最大にするような重み

係数ベクトルと分散の最尤推定値を $\tilde{\beta}, \tilde{\sigma}^2$ と表そう。NNモデルの作成にあたっては、代替的なネットワークパターンを想定した。その中で、符号条件を満足し、かつ尤度比が最大になるようなNNモデルを最終的に選択した。図－3には、最終的に選択したNNモデルを示している。

以上で推定した重み係数ベクトルの最尤推定値 $\tilde{\beta}$ を用いて事前分布を定義した。さらに、MH法により100,000個のパラメータサンプリングを実施した。その際、最尤推定値を初期パラメータサンプルと考え、 $\beta^{(0)} = \tilde{\beta}$ と設定した。図－4は、NNモデルの重み係数 β_{11} のサンプリング過程を表している。最尤推定値を初期サンプルとして用いているため、MH法によるパラメータサンプリングの過程において早い段階で定常過程に収束している。MH法によるパラメータサンプルとして10,000回以降のサンプルを有効サンプルとしてとりあげた。図－5は、MH法により求めたパラメータ β_{11} の事後周辺分布を示している。一方、分散 σ^2 に関しては、初期サンプルとして $\sigma^{2(0)} = \tilde{\sigma}^2$ を用いた。分散に関しても、サンプリング過程の早い段階で定常過程に収束している。図－6は、10,000回以降のサンプルを有効サンプルと考え、分散の事後周辺分布を求めた結果を表している。事前分布として無情報事前分布を仮定したにもかかわらず、事後周辺分布は対数正規分布に近い形を示している。つまり、事後周辺分布を推計する際、尤度関数が支配的な情報を与え、事前分布の影響はあまり残っていない。

表－2にベイズ推計により求めた重み係数、分散の期待値 $\tilde{\mu}(\beta_p)$ ($p = 1, \dots, s$), $\tilde{\mu}(\sigma^2)$, および分散 $\tilde{v}(\beta_p)$, $\tilde{v}(\sigma^2)$ を一括表示している。さらに、90%信頼区間、Geweke検定統計量を併記している。危険率2%での片側検定の場合、すべてのパラメータが収束することが分かる。

つぎに、各地点特性の重要度を評価するために、地点特性の影響度を定義しよう。実測サンプル k ($k = 1, \dots, M$)の土壌汚染浄化費用 $\hat{y}^k$ に対する地点特性 $\hat{x}_i^k$ の影響度 $R_i^k(\beta)$ をNNモデルを用いて

$$\begin{aligned} R_i^k(\hat{x}^k, \beta) &= \frac{\partial f(\hat{x}^k, \beta)}{\partial x_i^k} \\ &= \frac{\exp\{\alpha_0 + \sum_{j=1}^m \alpha_j F(\sum_{i=0}^n \beta_{ij} \hat{x}_i^k)\}}{\{\exp(\alpha_0 + \sum_{j=1}^m \alpha_j F(\sum_{i=0}^n \beta_{ij} \hat{x}_i^k) + 1\}^2} \\ &\quad \cdot \sum_{j=1}^m \left\{ \alpha_j \frac{\exp(\sum_{i=0}^n \beta_{ij} \hat{x}_i^k)}{\{\exp(\sum_{i=0}^n \beta_{ij} \hat{x}_i^k) + 1\}^2} \beta_{ij} \right\} \end{aligned} \quad (5.2)$$

と定義しよう．したがって，要素 x_i の影響度は

$$ER_i = \int_{\Theta/\sigma^2} \frac{R_i^k(\hat{\mathbf{x}}^k, \boldsymbol{\beta})}{M} \pi(\boldsymbol{\beta} | \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) d\boldsymbol{\beta} \quad (5.3)$$

と定義しよう．ただし， Θ/σ^2 は分散 σ^2 を除くパラメータ空間である．さらに， $\pi(\boldsymbol{\beta} | \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M)$ は，重み係数ベクトル $\boldsymbol{\beta}$ に関する事後周辺確率密度関数であり， $\pi(\boldsymbol{\beta} | \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) = \int_0^\infty \pi(\boldsymbol{\beta}, \sigma^2 | \hat{\boldsymbol{\xi}}^1, \dots, \hat{\boldsymbol{\xi}}^M) d\sigma^2$ と定義できる．

各地点特性の影響度を求めた結果を図－7に示している．同図より，表土の汚染範囲が土壌汚染浄化費用に最も大きな影響を及ぼすことが理解できる．また，重金属汚染と比較して，VOC汚染の影響度が大きい．VOC物質は拡散性が大きく，土壌汚染の範囲が大きくなるため，土壌汚染浄化費用が大きくなる．

5.3 土壌汚染浄化費用リスクのベイズ予測

34個の実測サンプルに対して，地点特性を入力情報として，BNNモデルを用いて土壌汚染浄化費用リスクを求めた．その際，MH法により，100,000個のパラメータサンプルを提案し，10,000個以降のパラメータサンプルを有効サンプルとして採用した．図－8は，実測サンプル($k=34$)の土壌汚染浄化費用リスクを求めた結果を示している．なお，実測サンプル($k=1$)の土壌汚染浄化費用リスクに関しては，すでに図－1に例示している．土壌汚染浄化費用モデルの確率項が対数正規分布に従うため，汚染浄化費用がNNモデルの推計結果より，相当程度過大になるリスクが存在する．本実測サンプルの場合，NNモデルによる推計結果と実績値の乖離はそれほど大きくない．しかし，危険水準5%のVaR値は実績値の10倍程度大きい値となっている．表－3には，各実測サンプルの土壌汚染浄化費用の実績値，NNモデルによる推定値，及び危険水準を5%に設定したVaR値も示している．同表に示すように，土壌汚染浄化費用の実績値は，すべて5%以下に収まっている．しかし，実績値とNNモデルによる最尤推計値が相当程度に乖離している実測サンプルもある．特に，表土調査の段階では，汚染物質が浸透した深度に関する情報が得られないため，簡易調査の結果のみでは土壌汚染浄化費用リスクが相当程度に存在することが理解できる．土地開発者は図－8に示すような土壌汚染浄化費用リスク情報に基づいて，土地開発の実施の有無や開発案の内容を評価することが求められる．土地所有者が，さらに精度の高い土壌汚染浄化費用の見積もりが必要であると判断した場合，ボーリング調査等が実施されることになる．

6 おわりに

土地開発の企画・計画を検討する構想段階では、土壤汚染の有無や土壤汚染浄化費用に関するリスク判断が必要となる。ボーリング調査等による詳細な土壤汚染調査には多額の費用が必要となることから、地籍調査、表土調査という簡易調査に基づいて、土壤汚染浄化費用を概算することが求められる。しかし、地盤条件が均質でない場合、簡易調査結果に基づいて見積もった土壤汚染浄化費用には、多くの推計誤差が存在すると考えることが妥当であろう。本研究では、BNNモデルを用いて、簡易調査結果より土壤汚染浄化費用の概算値とその変動リスクを評価する方法論を提案した。さらに、現実の土壤汚染処理費用データに基づいて、本研究で提案した方法論の有効性を実証的に検証した。

本研究で提案したBNNモデルは汎用性が高く、ベイズ予測モデルの核となるNNモデルを修正することにより、土壤汚染浄化費用リスクのみならず、多様なリスク評価問題に適用が可能である。しかし、土壤汚染浄化費用リスクの評価問題に絞っても、BNNモデルの有効性をより高めるために今後いくつかの課題が残されている。第1に、本研究では浄化工法として、掘削除去工法を想定した。浄化工法が異なれば、土壤汚染浄化費用リスクが異なる。今後は、浄化工法の多様性にも配慮できるような、BNNモデルを開発することが必要となる。第2に、本研究で得られた土壤汚染浄化費用リスクの情報を土地開発の意思決定に活用するための方法論を開発する必要がある。土地開発に関する意思決定には、土壤汚染浄化費用リスクだけでなく、マーケットリスク等の多様なリスクが関与する。土地所有者は、土地開発に関わる多様なリスクを総合的に判断し、土地開発の有無、開発の内容を決定したり、ボーリング調査等、詳細な土壤汚染調査を実施するか否かを決定することが求められる。土地開発、および土壤汚染調査を継続するか否かに関する意思決定問題は、リアルオプション問題として定式化が可能である。本研究で求めた費用リスクに関する情報は、リアルオプション問題を検討する際の重要な入力情報となり得る。第3に、事前分布、尤度関数の特定化誤差の問題がある。本研究では、NNモデルの観測誤差が対数正規分布に従うと仮定した。対数正規分布の仮定は、費用リスクの評価に重大な影響を及ぼすことになる。このような特定化誤差を回避するためには、関数形を特定化しないノンパラメトリックなBNNモデルの開発が必要である。第4に、土壤汚染浄化費用算定プロセスの標準

化が必要となる．現在のところ，利用可能な土壌汚染浄化費用データは，必ずしも標準化された方法で計上された費用データではない．このため，費用データの信頼性には改善すべき点が残されている．本研究で求めた土壌汚染浄化費用リスクは，土壌汚染に関連する保険料を算定するための有力な情報になりうる可能性がある．このような土壌汚染に関連する保険市場が発展するためには，土壌汚染浄化費用の算定方法の標準化が必要となる．最後に，実測サンプルの蓄積があげられる．わが国では，土壌汚染浄化対策が実施されて日も浅く，土壌汚染浄化費用に関するサンプルが蓄積されていないのが実情である．今後，新しい実測サンプルを獲得すれば，4.(5)で言及したような方法でBNNモデルをベイズ更新していく必要がある．

参考文献

- 1) 身崎成紀：土壌汚染リスク対策のポイント，2004
- 2) （株）三菱総合研究所 土壌汚染対策研究チーム：土壌汚染ビジネス大研究，2003
- 3) 森澤眞輔編著：土壌圏の管理技術，コロナ社，2002.
- 4) 林義隆，太田英将，國眼定：3次元空間解析による汚染地盤の調査数量と汚染量の信頼性について，地盤工学会関西支部地盤環境・計測シンポジウム，2003.
- 5) 石森洋行，阿武秀一，深川良一，極川哲也，佐藤千尋：土壌汚染分布の簡易的推定法の開発に関する基礎的研究，第11回地下水・土壌汚染とその防止対策に関する研究集会，2005.
- 6) Hertz, J., Krogh, A., and palmer, R.G.: *Introduction to the Theory of Neural Comutation*, Addison-Wesley, 1991.
- 7) Bishop, C.M.: *Neural Netorks for Pattern Recognition*, Oxford University Press, 1995.
- 8) Ripley, B.D.: *Pattern Recognition and Neural Networks*, Cambridge University Press, 1996.
- 9) Cybenko, G.: Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals, and Systems*, Vol.2, pp.303-324, 1989.
- 10) Hornik, K., Stinchcombe, M., and White, H.: Multilayer feedforward networks are universal approxiamtors, *Neural Networks*, Vol.2, pp. 356-366, 1989.
- 11) Box, G.E.P. and Tiao, G.C.: *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Company, 1973.
- 12) 繁枅算男：ベイズ統計入門，東京大学出版会，1985.
- 13) 鈴木雪夫，国友直人編：ベイズ統計学とその応用，東京大学出版会，1989.
- 14) 例えば、津田尚胤，貝戸清之，山本浩司，小林潔司：ワイブル劣化ハザードモデルのベイズ推計法，土木学会論文集(投稿中)。
- 15) Radford M.Neal： *Bayesian Learning for Neural Networks*, Springer Verlag, 1996
- 16) 和合肇：ベイズ計量経済分析，マルコフ連鎖モンテカルロ法とその応用，東洋経済新報社，2005.
- 17) Jeffreys, H.: *The Theory of Probability*, Oxford University Press, 1961.

- 18) Geman, S. and Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *Transactions on Pattern Analysis and Machine Intelligence*, Vol.6, pp.721-741, 1984.
- 19) Gelfand, A.E. and Smith, A.F.M.: Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association*, Vol.85, pp.398-409, 1990.
- 20) Gilks, W.R. and Wild, P.: Adaptive rejection sampling for Gibbs sampling, *Applied Statistics*, Vol.41, pp.337-348, 1992.
- 21) Chib, S.: Marginal likelihood from Gibbs output, *Journal of the American Statistical Association*, Vol.90, pp.1313-1321, 1995.
- 22) Han, C. and Carlin, B.P.: MCMC methods for computing Bayes factors, A comparative review, *Biometrika*, Vol.84, pp.711-732, 2000.
- 23) Geweke, J.: Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, in: Bernardo, J.M., Berger, J.M., Dawid, A.P., and Smith, A.F.M. (eds.) :*Bayesian Statistics 4*, pp.169-193, Oxford University Press, 1996.
- 24) Newey, W.K. and West, K.D.: A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, Vol.55, pp.703-708, 1987.

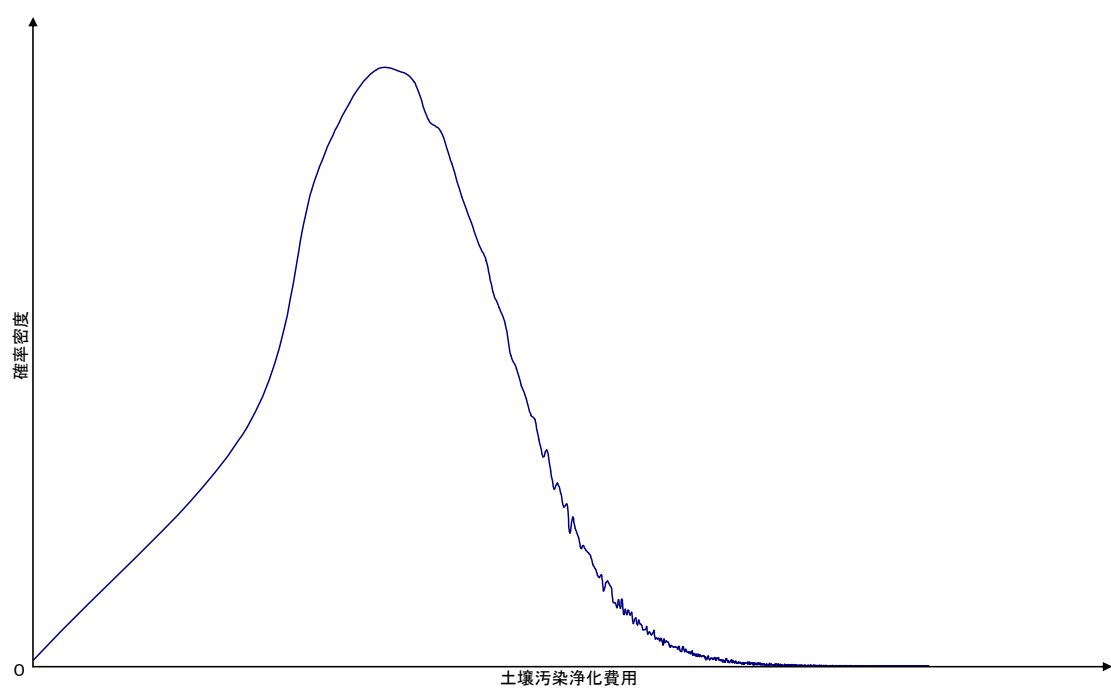


図 — 1 土 壌 汚 染 浄 化 費 用 リ ス ク

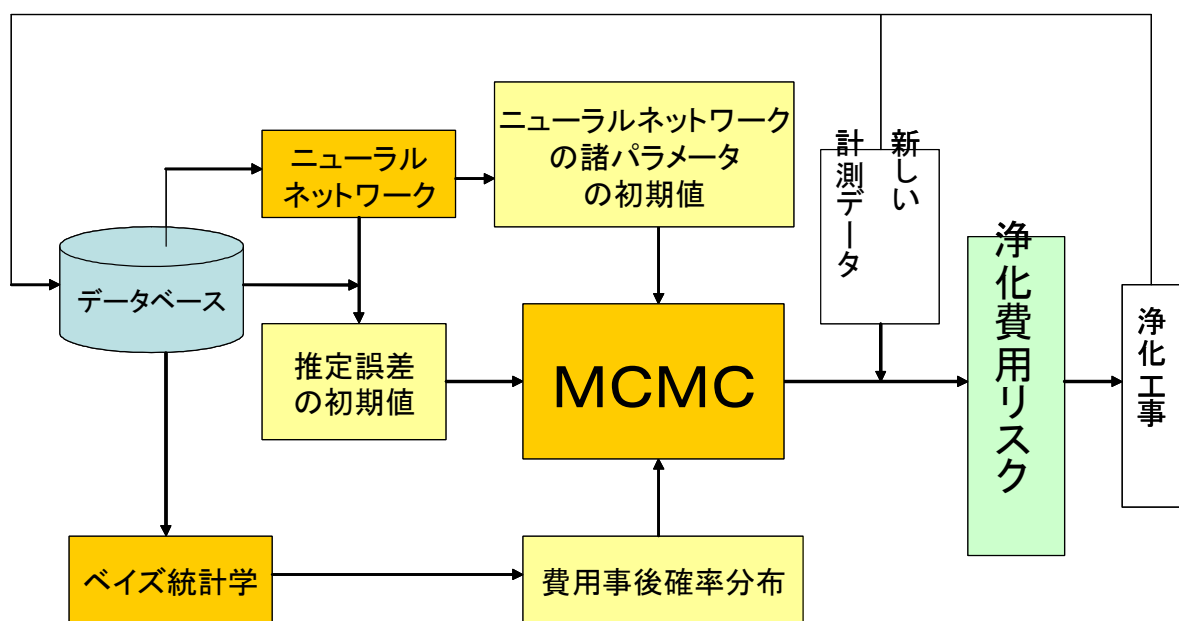


図 ー 2 土 土 壤 汚 染 浄 化 費 用 リ ス ク の 計 測 手 順

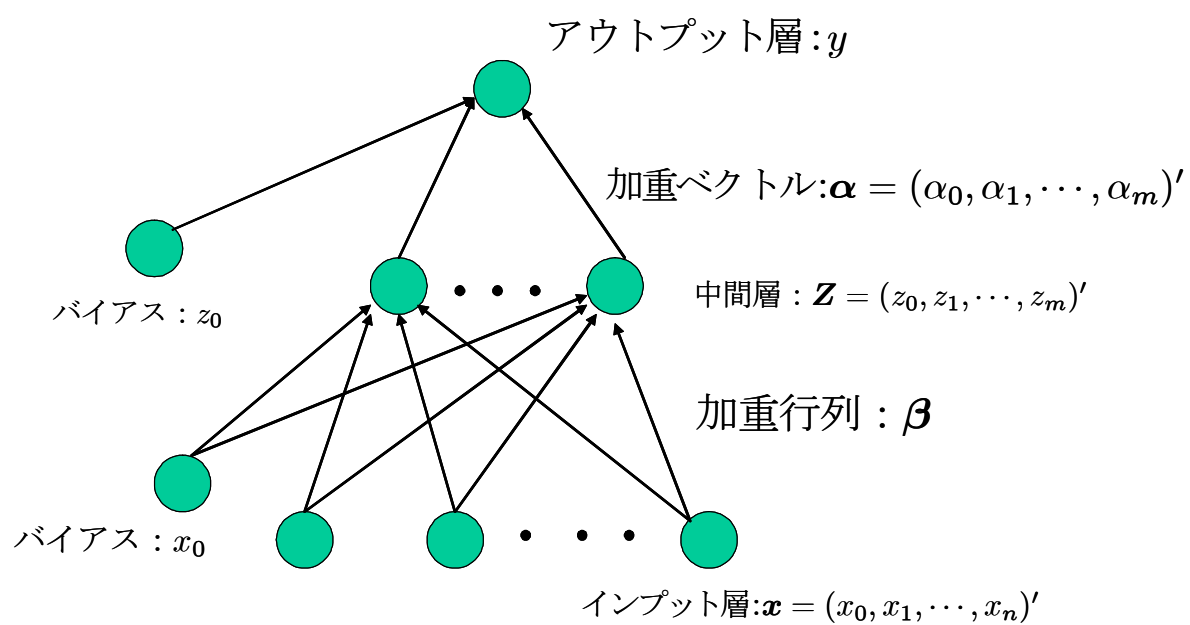


図 — 3 ニューラル ネットワーク

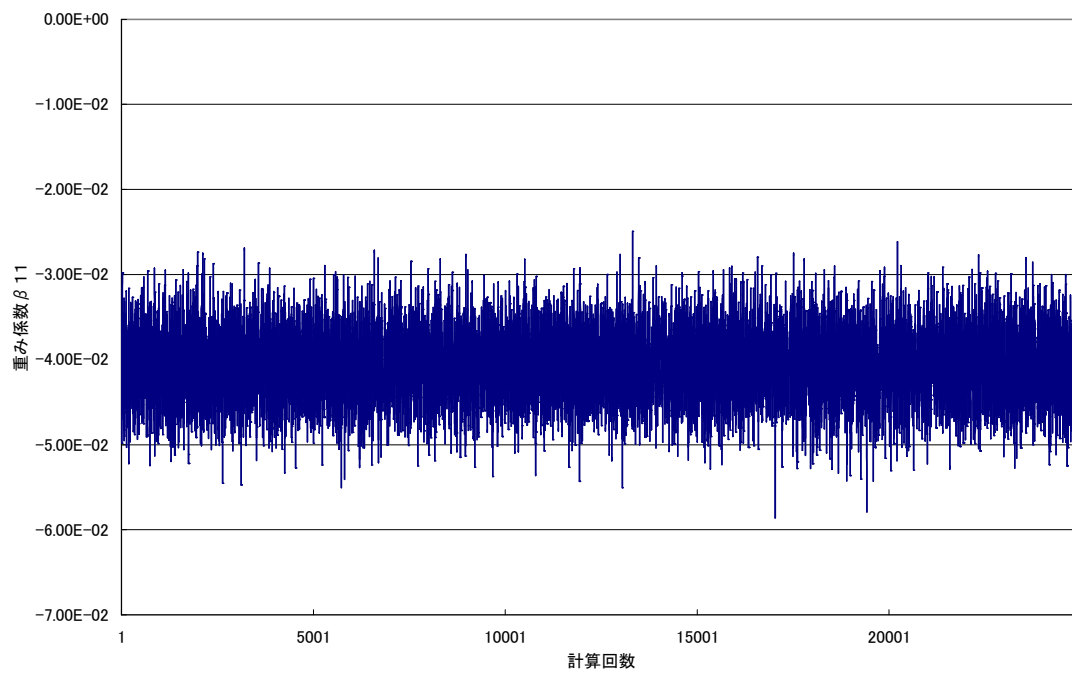


図 — 4 重み係数 β_{11} のサンプリング過程 (1)

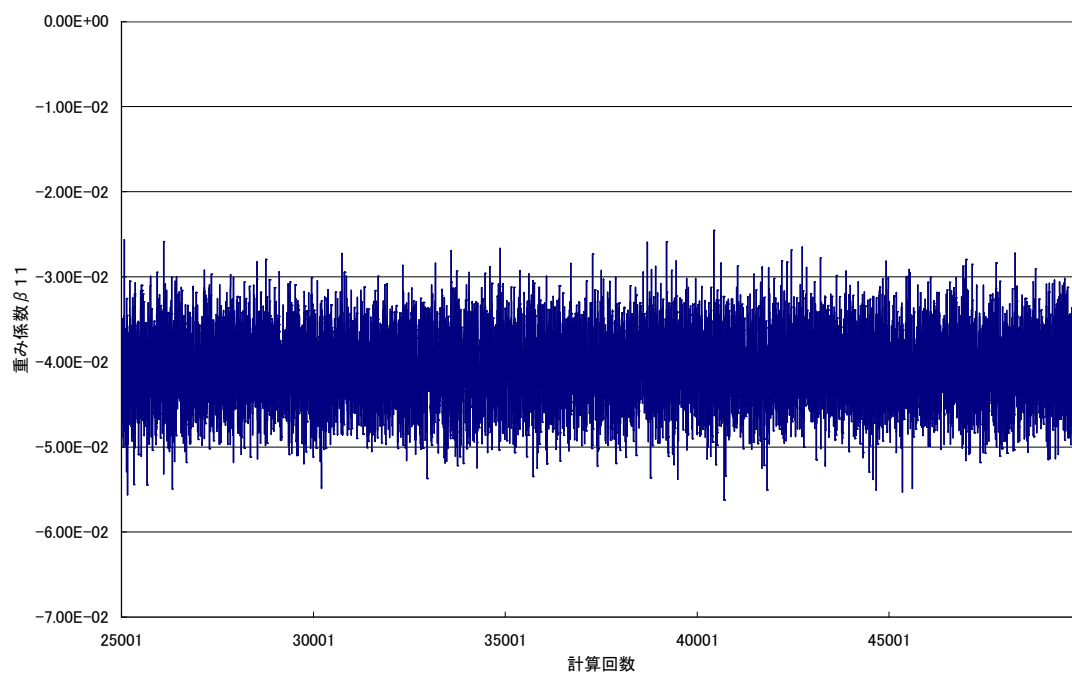


図 — 4 重み係数 β_{11} のサンプリング過程 (2)

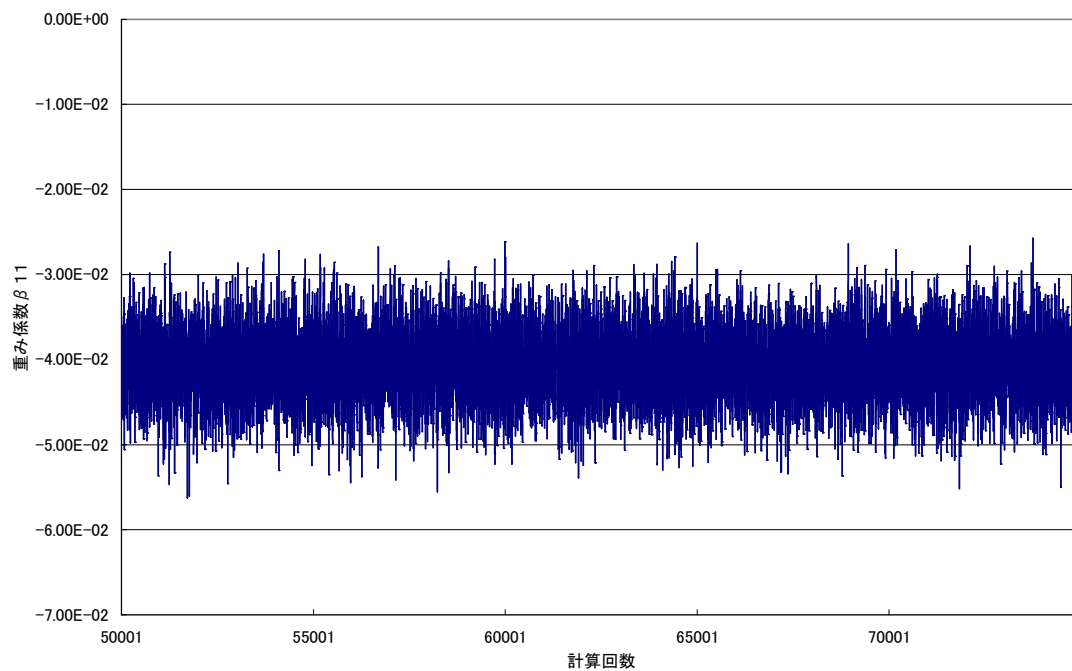


図 — 4 重み係数 β_{11} のサンプリング過程 (3)

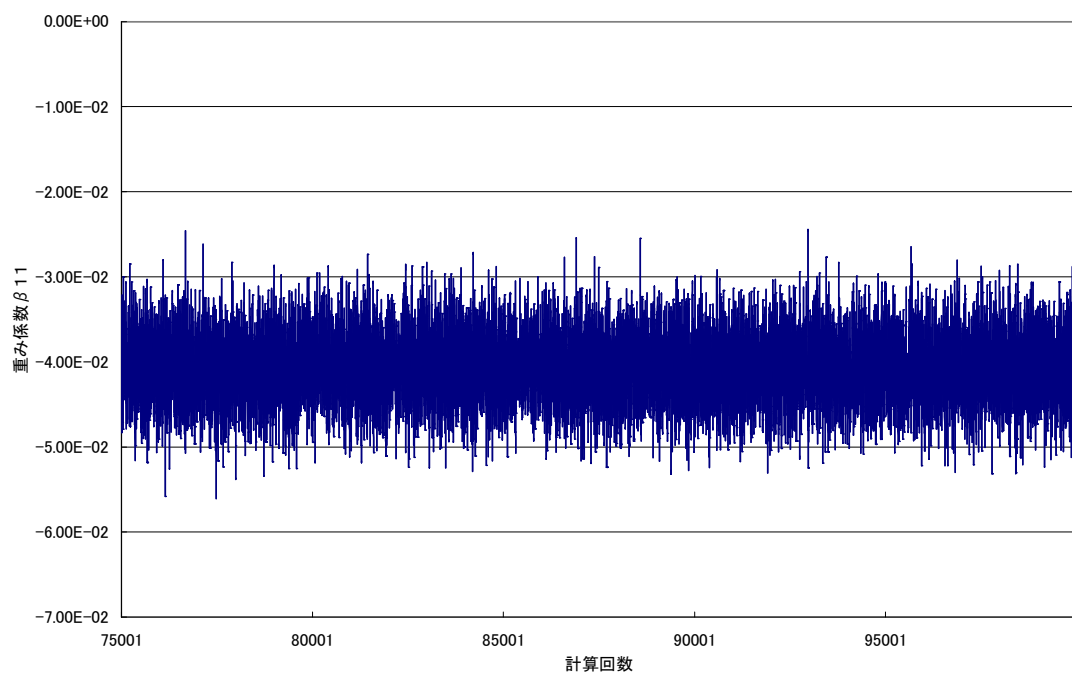


図 — 4 重み係数 β_{11} のサンプリング過程 (4)

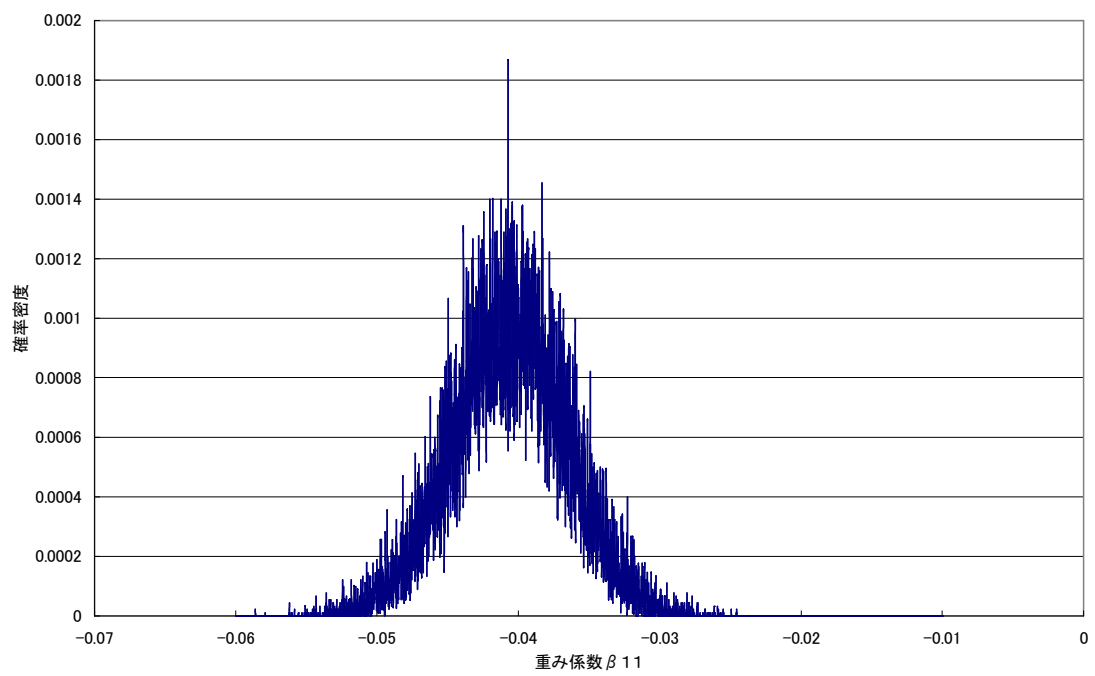


図 — 5 重み係数 β_{11} の事後周辺分布

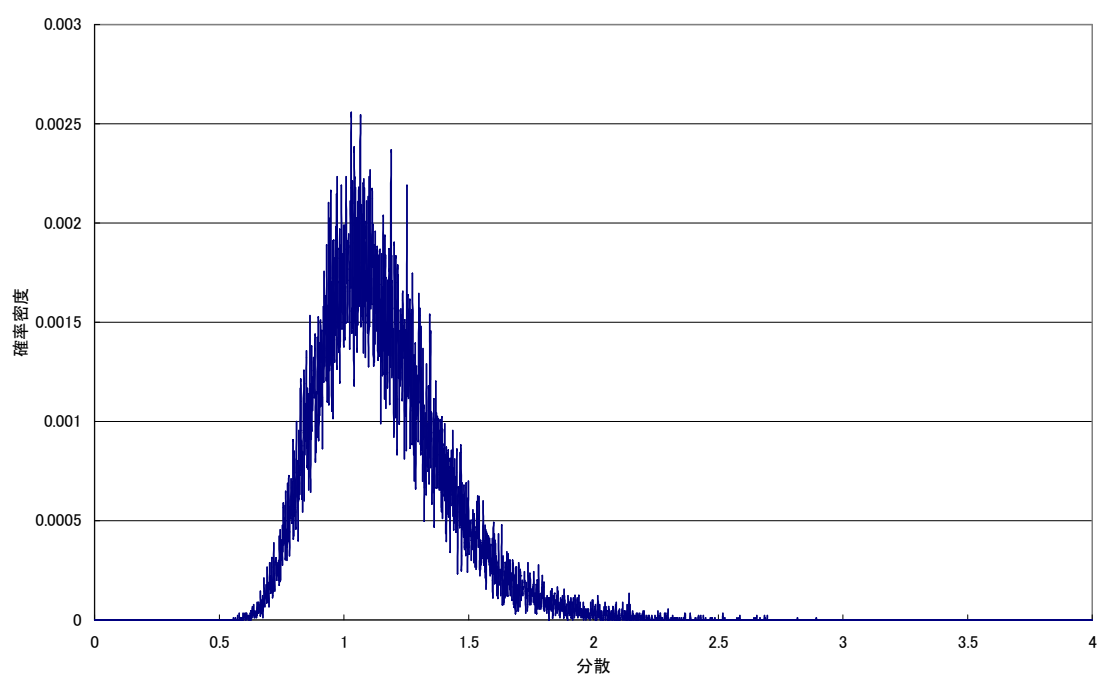


図 — 6 分散 σ^2 の事後周辺分布



図 — 7 地点特性の影響度

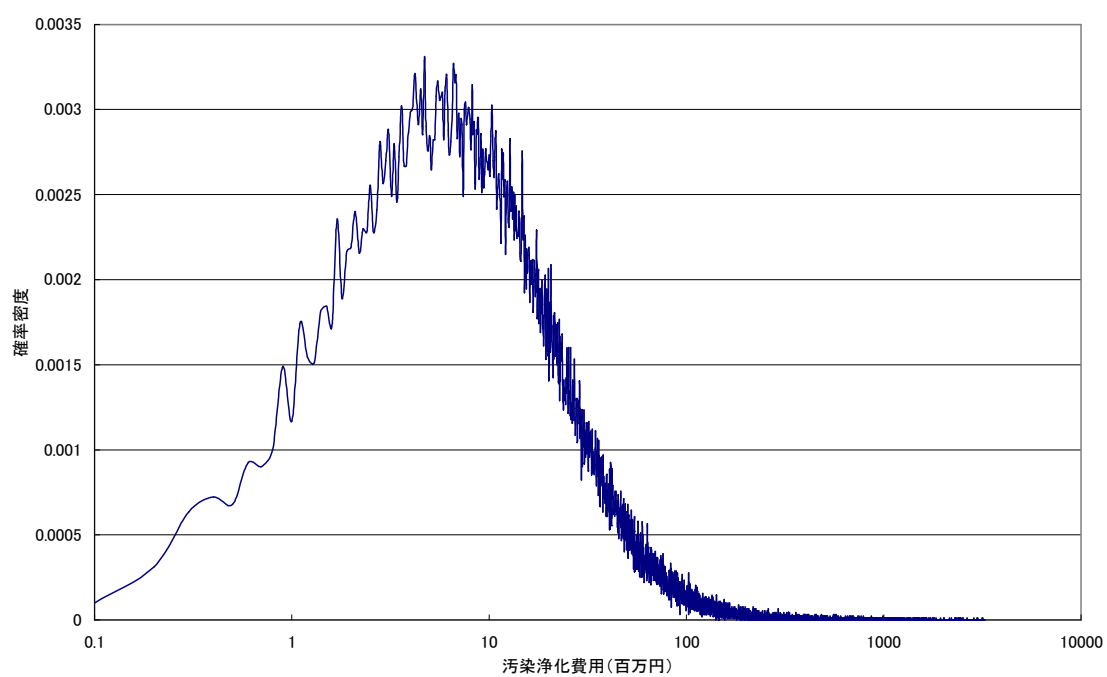


図 — 8 土 壌 汚 染 浄 化 費 用 リ ス ク （ 実 費 9 0 0 0 万 円 ）

表 — 1 地 点 特 性

変数	情報の内容
用途地域	住宅, 商業, 工業
地盤強度	軟弱, 普通, 強固
土地利用	クリーニング・GS(面積1000m2以下), 工場・廃材 置場・資材置場, 写真・印刷業・メッキ, 病院, 大 学, 住宅・店舗・事務所, 裸地・公園
敷地面積	単位: 平方メートル
汚染物質	重金属, VOC
汚染の範囲	単位: 平方メートル
汚染物濃度	単位: mg/l

表 — 2 B N N モデルの推計結果

	平均	分散	90%信頼区間		Geweke検定統計量
重み係数 β 1 $\rightarrow$ 1	-4.06E-02	1.66E-05	-4.74E-02	-3.40E-02	7.48E-02
重み係数 β 2 $\rightarrow$ 1	-8.79E-01	7.83E-03	-1.03E+00	-7.34E-01	5.40E-02
重み係数 β 3 $\rightarrow$ 1	7.22E-01	5.37E-03	6.01E-01	8.41E-01	6.52E-01
重み係数 β 4 $\rightarrow$ 1	-1.39E+00	1.96E-02	-1.62E+00	-1.16E+00	2.81E-02
重み係数 β 5 $\rightarrow$ 1	-1.46E+00	2.21E-02	-1.71E+00	-1.22E+00	4.91E-01
重み係数 β 6 $\rightarrow$ 1	-5.76E-02	3.32E-05	-6.72E-02	-4.83E-02	1.89E+00
重み係数 β 7 $\rightarrow$ 1	-1.07E+00	1.16E-02	-1.25E+00	-8.93E-01	5.07E-01
重み係数 β 8 $\rightarrow$ 1	2.95E-01	8.62E-04	2.47E-01	3.42E-01	4.18E-01
重み係数 β 9 $\rightarrow$ 1	-1.71E-01	2.91E-04	-1.99E-01	-1.43E-01	6.38E-01
重み係数 β 10 $\rightarrow$ 1	1.95E+00	3.70E-02	1.64E+00	2.26E+00	8.70E-01
重み係数 β 11 $\rightarrow$ 1	-4.03E-02	1.63E-05	-4.70E-02	-3.37E-02	2.52E-01
重み係数 β 12 $\rightarrow$ 1	-2.09E+00	4.53E-02	-2.44E+00	-1.75E+00	1.64E-01
重み係数 β 13 $\rightarrow$ 1	-1.15E+00	1.33E-02	-1.35E+00	-9.66E-01	1.11E+00
重み係数 β 14 $\rightarrow$ 1	2.88E+00	8.40E-02	2.40E+00	3.35E+00	3.38E-01
重み係数 β 15 $\rightarrow$ 1	1.12E+00	1.24E-02	9.39E-01	1.31E+00	7.75E-01
重み係数 β 16 $\rightarrow$ 1	1.02E+00	1.08E-02	8.46E-01	1.19E+00	1.99E+00
重み係数 β 17 $\rightarrow$ 1	3.04E+00	9.33E-02	2.53E+00	3.54E+00	3.52E-01
重み係数 β 18 $\rightarrow$ 1	1.53E+00	2.33E-02	1.27E+00	1.77E+00	5.84E-01
重み係数 β 19 $\rightarrow$ 1	4.95E-01	2.44E-03	4.14E-01	5.76E-01	2.03E-01
重み係数 β 4 $\rightarrow$ 2	-8.38E-01	6.98E-03	-9.76E-01	-7.01E-01	6.72E-01
重み係数 β 5 $\rightarrow$ 2	2.52E-01	6.35E-04	2.10E-01	2.93E-01	2.63E-01
重み係数 β 6 $\rightarrow$ 2	-1.24E+00	1.55E-02	-1.45E+00	-1.04E+00	9.45E-01
重み係数 β 15 $\rightarrow$ 2	-1.44E+00	2.07E-02	-1.68E+00	-1.21E+00	2.23E-01
重み係数 β 16 $\rightarrow$ 2	-1.28E+00	1.69E-02	-1.50E+00	-1.07E+00	2.38E-01
重み係数 β 18 $\rightarrow$ 2	-1.48E+00	2.24E-02	-1.73E+00	-1.24E+00	6.25E-01
重み係数 β 19 $\rightarrow$ 2	3.22E-01	5.33E-04	2.83E-01	3.59E-01	2.94E-01
重み係数 α 1 $\rightarrow$ 3	6.49E+00	4.55E-01	5.40E+00	7.60E+00	2.99E-01
重み係数 α 2 $\rightarrow$ 3	-2.29E+00	2.67E-02	-2.56E+00	-2.02E+00	1.01E+00
バイアス項係数 β 0 $\rightarrow$ 1	-2.45E+00	6.72E-02	-2.87E+00	-2.01E+00	7.37E-01
バイアス項係数 β 0 $\rightarrow$ 2	6.38E-01	2.06E-03	5.64E-01	7.13E-01	9.39E-01
バイアス項係数 α 0 $\rightarrow$ 3	-3.01E+00	6.24E-02	-3.42E+00	-2.60E+00	2.02E+00
分散	1.15E+00	6.21E-02	8.11E-01	1.61E+00	1.02E+00

表 — 3 ベ イ ズ 予 測 結 果

サンプル	浄化費用の実績値 (百万円)	NNモデル推定値 (百万円)	VaR(5%) (百万円)
1	21	22.94	141.4
2	120	51.27	208.9
3	4	28.3	152.3
4	117	242.06	516.9
5	25	39.57	204.4
6	105	314.37	533.3
7	600	784.07	1795.1
8	12	60.06	339.0
9	90	21.71	100.4

謝 辞

本論文を結ぶにあたり，本研究の遂行に際して，大変多忙な中，終始適切な御指導と御助言，そして論文の緻密な修正，暖かい励ましの言葉をいただきました京都大学大学院工学研究科の小林潔司教授に心から感謝申し上げます。また，京都大学防災研究所の多々納裕一教授には，本論文の審査，修正に際し，貴重な御助言，御示唆をいただきました。厚く御礼申し上げます。京都大学大学院工学研究科松島格也助教授にはご多忙の中，有益な御指導と御教授を頂き，心より感謝の意を表します。京都大学大学院工学研究科の大西正光助手には基礎的な素養や研究への姿勢をお教え頂きました。深甚なる感謝の意を表します。そして，計画マネジメント論研究室の諸兄には，本研究を取りまとめる上で多大な御協力を頂きました。ここに深く感謝する次第です。最後になりましたが，計画マネジメント論研究室の津田尚胤氏より有意義な議論を重ねながら研究を深め，またその合間には同氏より基礎知識をお教え頂き，心から感謝の意を表します。