

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 4 日現在

機関番号：17102

研究種目：基盤研究(B)

研究期間：2011～2014

課題番号：23300106

研究課題名(和文) 時空間現象データの統計モデリングと当該現象の定量的把握の研究

研究課題名(英文) Statistical modeling of spatio-temporal data and quantitative grasp of phenomena

研究代表者

西井 龍映(Nishii, Ryuei)

九州大学・マス・フォア・インダストリ研究所・教授

研究者番号：40127684

交付決定額(研究期間全体)：(直接経費) 8,900,000円

研究成果の概要(和文)：時空間現象から得られる種々のデータに対する統計モデルの開発と評価を行い、当該現象の特徴を把握し、モデルの統計科学的性質を研究した。具体的な課題は次の通りである。

1) 森林被覆率を人口密度や起伏量によって説明する時間的空間的依存性を考慮した回帰モデルを提案した。2) 太陽風の種々の物理量に関する時系列観測データを用いて磁気嵐指数を予測するモデルを提案した。3) 地震頻度と太陽風との関連を考察し、マグニチュード4以下の全球の地震頻度は、太陽風の影響があることを統計的に示した。4) 確率微分方程式で記述されるモデルの母数推定について、漸近理論を考察した。

研究成果の概要(英文)：The aim of our research was to promote statistical modeling for spatio-temporal data acquired by actual phenomena, and grasp its characteristics through the modeling. Also we studied statistical properties of the models. The concrete results are as follows.

1) We considered deforestation models where forest-covering-ratio is regressed by human population density and relief energy with spatial and temporal dependency. 2) Auto-regressive models with exogenous variables are extended through generalized linear models, and applied to actual data. Then, magnetic storm time index is accurately predicted by using physical measurements of solar activities. 3) Frequencies of weak earthquakes were shown to be affected by the solar wind. 4) Asymptotic theory for the parameter estimation of time series models described by stochastic differential equations was developed.

研究分野：統計学

キーワード：時空間モデリング 時系列解析 環境統計 ノンパラメトリック解析 森林減少 宇宙天気 確率微分方程式 モデル選択

1. 研究開始当初の背景

計測機器・センサの発展，データ転送の技術の進化，ネットワークの広域化・高速化により，従来は得られなかった高次元・高頻度の巨大な時空間データが蓄積されるようになってきた．しかし無意味な情報やノイズも多く含まれているため，データの量に比例して有益な情報が増えるわけではない．ここでは地球環境を観測した時空間データや企業の製品開発現場から得られる時系列・画像等の時空間データに対し，その時間的空間的ダイナミクスを統計モデルに取り込み，そのモデルを通して現象の特徴を把握し，現象を解明する．

2. 研究の目的

時空間現象の解析に関し，次の具体的な3課題を研究することを目的とした．

1) リモートセンシング等により得られた地球環境の時空間データに対する統計モデルの構築・推定と評価方法を開発し，環境の空間的及び時間的ダイナミクスを統計モデルにより検証する．

2) 企業の開発現場から得られるデータ (多次元高頻度データ，画像・感性データを含む) に対する統計モデルを開発し，その統計科学的性質を研究する．

3) 各分担者は，それぞれの専門分野での時空間現象に対し，統計モデリング及びその統計的性質を解明する研究を行う．

3. 研究の方法

各課題について，次の研究方法を用いた．

1) 小領域での森林面積比率を人口密度と起伏量で説明する非線形回帰モデルを考察した．特に森林比率が 0, 1 および区間(0,1)の値をとる場合を 3 群の判別問題，また比率が (0,1) に含まれるときは回帰モデルを考えると，統一的なモデリングを考察した．また時間・空間に関する依存性についても考察した．

2) 企業の開発現場での時系列データの予測方法について自己回帰項をもつ回帰モデル (ARX モデル) を利用し，モデル選択基準を導出した．さらにその一般化加法モデル (GAM) により，実データに適用した．一方高頻度データの確率微分方程式等の母数推定に関する漸近理論を考察した．

3) 方向統計学のモデル開発や基礎理論の構築，回帰モデルのスージングとその漸近理論に関する研究を行った．

4. 研究成果

(1) 森林被覆率を人口密度と起伏量で説明する回帰モデルにおいて，自然三次スプラインで記述される柔軟な平均構造を持つ非線形回帰モデルを考察した．これは従来のパラメトリックな平均構造で記述されるモデルを大きく改良した．また同様な結果はサポートベクター回帰 (SVR) でも得られることがわかった．

(2) 森林被覆率を目的変数とするメッシュで観測した格子状のデータに関する回帰モデルについて，近傍画素からの影響をモデル化した空間依存性をもつ 0-1 インフレート分布を提案した．また巾変換やスプライン関数により，説明変数の影響の非線形性をモデル化した．機械学習によるアプローチとして，3 群判別をサポートベクターマシン (SVM)，空間依存性を持つ回帰モデルについては SVR が有用であることを確認した．

(3) 強い太陽風が地磁気に影響を及ぼし，地震誘発のトリガとなっているとの仮説が 20 世紀終わりに提唱されたが，明確な結論が得られていない．この仮説を正規分布，ポアソン回帰や負の 2 項分布に基づく一般化線形モデルで検証した．これにより太陽風が弱い地震のトリガになっていることや分散構造にもその影響があることがわかった．

(4) 磁気嵐指数 Dst を自己回帰および太陽風の物理量から予測する ARX モデルを決定した．また分散も変化する GAM モデルを考察し，新しいモデル評価基準を提案した．さらに Dst は裾の重い分布に従うことから，誤差を t-分布に従うとし，スケールや自由度も GAM で記述される一般的なモデルの中から最適モデルを選び，予測精度の高いモデルを得た．

(5) 自己回帰成分を持つ時系列データの回帰モデルで，自己回帰の次数，説明変数や時間遅れ次数を選択する問題について，重み付き最小 2 乗法による母数推定による GIC 基準を求め，車両運転装置に応用する装置を企業と共同開発し，特許を公開した．また実験計画において，モデル推定に有用な追加標本を高速に選択する手法を提案した．提案手法を自動車の剛性を一定値以上に保ち，かつ軽量の車体を設計することに応用し，その有効性を確認した．

(6) 地表面の超高次元分光画像に基づき、森林・畑等の土地カテゴリーの被覆割合を推定する問題 (unmixing) を考察した。ディリクレ分布を被覆割合の事前分布とする空間依存性を持つベイズモデルによる手法を提案し、実データで高性能であることを示した。

(7) 任意次元の確率微分方程式モデルからの高頻度データに基づいた正規型疑似尤度推定について、多項式型大偏差不等式を介し、推定量のモーメントの収束まで導出した。

- ・ 高頻度観測される Normal inverse Gaussian 過程の局所漸近正規性を導出した。
- ・ 夜間・昼休みの経済高頻度データへの影響を加味したボラティリティの推定手法を定式化し、実証分析でその安定性を示した。
- ・ Tempered stable Ornstein-Uhlenbeck 過程の遷移確率構造を導出した。また大規模従属データ解析のための漸近推測理論および関連する確率過程や統計的確率場の極限定理や分布論を、R パッケージを介した計算機上の実装環境の構築と並行して研究した。
- ・ Levy 過程モデルからの大規模高頻度データに基づいた統計推測理論に関する総合報告をまとめた。

(8) 非線形関数で構成した基底関数のそれぞれのウェイトに事前分布を設定して、予測に有効なスパースなモデルを構築する関連ベクターマシンについて、様々な事前情報をモデルに取り込むベイズアプローチによる非線形回帰予測分布モデルの構成法、モデルの推定法と評価法について研究した。また正則化法の標準ツールである LASSO に対し AIC 元来の定義に基づいた情報量規準を提案した。また LASSO より性能が高いといわれる非凹罰則付き最尤法に対応する同様の情報量規準を新たに導き、性能を数値実験で確認した。

(9) 離散型非対称分布としての一般化非心負の二項分布の構成と分布の性質について調べた。また、時空間現象としての海面水位の予測のために角度データである風向を取り入れたモデル化、および軸分布の構成について研究した。さらに角度を含むデータの解析とモデル化を扱う方向統計学の研究を行った。理論的側面に関しては、ハート型分布の拡張を行うとともに非斉次隠れマルコフモデルを提案し、応用面に関しては、環境統計学の立場から地震データや気象データを解析した。

(10) カーネル型確率点推定量に関するエッジワース展開を求め、その有効性を示した。さらにカーネル型確率密度関数推定量の高次のエッジワース展開を具体的な形で求めた。また順位検定統計量の連続化を提案し、離散統計量の有意確率についての問題点が

解消できることを示した。またカーネル型推定を利用して順位統計量を連続化した検定統計量の高次漸近理論構築し、エッジワース展開を求めた。さらにハザード関数のノンパラメトリック推定量についての漸近表現を求め、推測精度を改善した

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[1] R. Nishii (2014). Regression analysis and its development. A Mathematical Approach to Research Problems of Science and Technology- Theoretical Basis and Developments in Mathematical Modeling. Mathematics for Industry 5, R. Nishii et al. (eds.), Springer, 249-262. (図書編集と寄稿)

[2] P. Qin, T. Yamasaki and R. Nishii (2014). Statistical detection of the influence of solar activities to weak earthquakes. Pacific Journal of Math-for-Industry 6:6.

[3] Y. Ninomiya (2014). Change-point model selection via AIC. To appear in Annals of the Institute of Statistical Mathematics.

[4] D. Kim, S. Kawano and Y. Ninomiya (2014). Adaptive basis expansion via L1 trend filtering. Computational Statistics 29, 1005-1023.

[5] H. Masuda (2015). Parametric estimation of Levy processes. Levy Matters IV, Lecture Notes in Mathematics 2128, Springer, 179-282.

[6] R. Nishii and S. Tanaka (2013). Modeling and inference of forest coverage ratio using zero-one inflated distributions with spatial dependence. Environmental and Ecological Statistics 20 (2), 315-336.

[7] Y. Maesono and S.I. Penev (2013). Improved confidence intervals for quantiles. Annals of the Institute of Statistical Mathematics 65, 167-189.

[8] H. Masuda (2013). Convergence of Gaussian quasi-likelihood random fields for ergodic Levy driven SDE observed at high frequency. Annals of Statistics 41, 1593-1641.

[9] H. Masuda (2013). Asymptotics for functionals of self-normalized residuals of discretely observed stochastic processes. *Stochastic Processes and their Applications* 123(7), 2752-2778.

[10] P. Qin, R. Nishii and Z. Yang (2012). Selection of NARX models estimated using weighted least squares method via GIC-based method and L_1 -norm regularization methods. *Nonlinear Dynamics* 70, 1831-1846.

〔雑誌論文〕(計 32 件)

〔学会発表〕(計 64 件)

〔図書〕(計 3 件)

〔産業財産権〕

取得状況 (計 1 件)

名称：車両用運転支援装置
発明者：西井 龍映, 秦 攀, 新部 忠幸,
中本 尊元, 楠本 信平
権利者：九州大学, マツダ株式会社
種類：特許権
番号：特願 2011-158888
出願年月日：2011 年 07 月 20 日
取得年月日：2013 年 02 月 04 日
国内外の別：国内

〔その他〕

ホームページ等

下記に公開予定

(田中分担者が島根大から移動したため)

<http://tathagata.tanakalab.hue.ac.jp/op/enProgs/index.html>

5. 研究組織

(1) 研究代表者

西井 龍映 (NISHII RYUEI)
九州大学・マス・フォア・インダストリ
研究所・教授
研究者番号：40127684

(2) 研究分担者

二宮 嘉行 (NINOMIYA YOSHIYUKI)
九州大学・マス・フォア・インダストリ
研究所・准教授
研究者番号：50343330

増田 弘毅 (MASUDA HIROKI)
九州大学・マス・フォア・インダストリ
研究所・准教授
研究者番号：10380669

前園 宜彦 (YOSHIHIKO MAESONO)
九州大学・大学院数理学研究院・教授
研究者番号：30173701

田中 章司郎 (SHOJIRO TANAKA)
島根大学・大学院総合理工学研究科・
教授
研究者番号：00197427

清水 邦夫 (SHIMIZU KUNIO)
慶應義塾大学・名誉教授
研究者番号：60110946

秦 攀 (QIN PAN)
(H25 年末まで)
九州大学・マス・フォア・インダストリ
研究所・学術研究員
研究者番号：40532718

小西 貞則 (KONISHI SADANORI)
(H24 年度まで)
中央大学・理工学部・教授
研究者番号：40090550

坂田 年男 (SAKATA TOSHIO)
(H24 年度まで)
九州大学・大学院芸術工学研究院・教授
研究者番号：20117352