

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-191630

(P2019-191630A)

(43) 公開日 令和1年10月31日(2019.10.31)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G 1 6 Z 99/00</b> (2019.01)	G O 6 F 19/00 1 1 O	5 L O 4 9
<b>G O 6 F 16/00</b> (2019.01)	G O 6 F 17/30 2 2 O Z	

審査請求 未請求 請求項の数 10 O L (全 14 頁)

(21) 出願番号	特願2018-79640 (P2018-79640)	(71) 出願人	504132272
(22) 出願日	平成30年4月18日 (2018. 4. 18)		国立大学法人京都大学
	(出願人による申告) 平成28年度、国立研究開発法人科学技術振興機構、戦略的創造研究推進事業 研究題目「野外の生物群集ネットワークを利用した植物の動態予測」に係る委託研究、産業力強化法第19条の適用を受ける特許出願	(74) 代理人	100111567
			弁理士 坂本 寛
		(72) 発明者	潮 雅之
			京都府京都市左京区吉田本町36番地1
			国立大学法人京都大学内
		Fターム(参考)	5L049 DD02

(54) 【発明の名称】 解析装置、解析方法、及びコンピュータプログラム

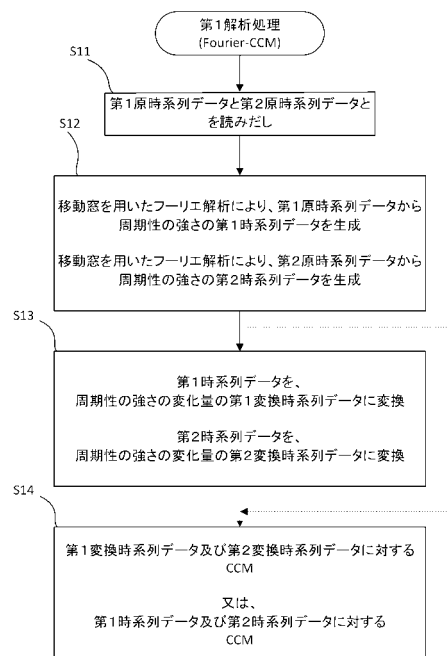
(57) 【要約】 (修正有)

【課題】 周期性を持つ時系列データに適した因果関係解析の手法を提供する。

【解決手段】 解析処理は、第1原時系列データから、第1原時系列データにおける周期性の強さの第1時系列データを生成し、第2原時系列データから、第2原時系列データにおける周期性の強さの第2時系列データを生成するステップS12と、第1時系列データ及び第2時系列データに基づいて、因果関係解析をすることを含む第1解析処理及び第2解析処理を行うステップS14と、を含む。

【選択図】 図1

図1



**【特許請求の範囲】****【請求項 1】**

第 1 原時系列データから、前記第 1 原時系列データにおける周期性の強さの第 1 時系列データを生成し、

第 2 原時系列データから、前記第 2 原時系列データにおける周期性の強さの第 2 時系列データを生成し、

前記第 1 時系列データ及び前記第 2 時系列データに基づいて、因果関係解析をすることを含む第 1 解析処理を実行する解析装置。

**【請求項 2】**

前記第 1 解析処理は、

前記第 1 時系列データを、前記第 1 時系列データにおける周期性の強さの変化量の時系列データである第 1 変換時系列データに変換し、

前記第 2 時系列データを、前記第 2 時系列データにおける周期性の強さの変化量の時系列データである第 2 変換時系列データに変換する

ことを更に含み、

前記因果関係解析は、前記第 1 変換時系列データ及び前記第 2 変換時系列データに対して行われる

請求項 1 に記載の装置。

**【請求項 3】**

前記第 1 時系列データは、前記第 1 原時系列データに対して移動窓を用いたスペクトル解析をすることで生成され、

前記第 2 時系列データは、前記第 2 原時系列データに対して移動窓を用いたスペクトル解析をすることで生成される

請求項 1 又は 2 に記載の装置。

**【請求項 4】**

前記第 1 原時系列データ及び前記第 2 原時系列データそれぞれにおける周期性の存在を識別する処理を更に実行する

請求項 1 から 3 のいずれか 1 項に記載の装置。

**【請求項 5】**

前記識別する処理の結果は、前記第 1 解析処理の実行の要否の判断に用いられる

請求項 4 に記載の装置。

**【請求項 6】**

前記第 1 解析処理は、前記第 1 原時系列データの周期性と前記第 2 原時系列データの周期性とが同調している場合に実行される

請求項 1 から 5 のいずれか 1 項に記載の装置。

**【請求項 7】**

前記第 1 原時系列データ及び前記第 2 原時系列データに対して前記因果関係解析をすることを更に実行する

請求項 1 から 6 のいずれか 1 項に記載の装置。

**【請求項 8】**

前記第 1 解析処理の結果及び前記第 2 解析処理の結果を出力する

請求項 7 に記載の装置。

**【請求項 9】**

第 1 原時系列データから、前記第 1 原時系列データにおける周期性の強さの第 1 時系列データを生成し、

第 2 原時系列データから、前記第 2 原時系列データにおける周期性の強さの第 2 時系列データを生成し、

前記第 1 時系列データ及び前記第 2 時系列データに基づいて、因果関係解析をすることを含む解析方法。

**【請求項 10】**

10

20

30

40

50

コンピュータに第1解析処理を実行させるコンピュータプログラムであって、前記第1解析処理は、

第1原時系列データから、前記第1原時系列データにおける周期性の強さの第1時系列データを生成し、

第2原時系列データから、前記第2原時系列データにおける周期性の強さの第2時系列データを生成し、

前記第1時系列データ及び前記第2時系列データに基づいて、因果関係解析をすることを含むコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本開示は、時系列データの因果関係の解析に関する。

【背景技術】

【0002】

時系列データは、ある現象を定期的に観察・測定し得られるデータである。時系列データを解析するためのデータ解析手法を時系列解析と言う。時系列解析において、時系列データ間の因果関係を解析する手法が注目されている。因果関係の解析は、例えば、時系列データ同士の因果関係の強さ又は方向を明らかにすることである。因果関係解析の手法としては、例えば、非特許文献1に示すConvergent Cross Mapping (CCM)、非特許文献2に示すTransfer Entropy (TE)、及び非特許文献3に示す状態空間モデルを用いた時系列解析などがある。

20

【先行技術文献】

【非特許文献】

【0003】

【非特許文献1】Sugihara et al., "Detecting causality in complex ecosystems", Science 338:496-500 (2012)

【非特許文献2】Thomas Schreiber, "Measuring Information Transfer", PHYSICAL REVIEW LETTERS VOLUME 85, NUMBER 2, 10 JULY 2000

【非特許文献3】深谷 肇一、「状態空間モデルによる時系列解析とその生態学への応用」、日本生態学会誌 66:375-389(2016)

30

【発明の概要】

【0004】

従来因果関係解析手法では、周期性を持つ時系列データ間の因果関係を正しく検出できないことがある。これは、周期性を持つ時系列データ間では、因果がなくても、時系列データの振動パターンが似通る同調性のためである。時系列データ同士が同調していると、実際には因果がない時系列データ間において、因果関係が誤検出されることがある。

【0005】

世の中の多くの時系列データは、周期性(例えば24時間周期)を持つ。例えば、1時間毎に野外で気温や光強度を測定すれば明瞭な24時間周期を観察できる。また、ヒトを含む生物全般で代謝活性・遺伝子発現量などで24時間の周期性が確認されている。生物における24時間の周期性は、概日リズムと呼ばれている。しかし、このような24時間周期を持つ全ての時系列データ間に因果があるわけではなく、時系列データの振動パターンが単に24時間周期で同調しているにすぎないことが多い。

40

【0006】

また、周期性を持つ時系列データは、森林の季節性、人間の行動パターン、気象などでも普遍的にみられる。例えば、1ヶ月ごとに日本の森林でドングリの結実をモニタリングすれば明瞭な1年周期が観察される。このように、世の中には、周期性を持つ時系列データが普遍的に存在する。したがって、周期性を持つ時系列データに適した因果関係解析の手法が望まれる。

【0007】

50

本開示において、因果関係解析は、周期性の強さの時系列データに基づいてなされる。更なる詳細は、後述の実施形態として説明される。

【図面の簡単な説明】

【0008】

【図1】図1は、第1解析処理のフローチャートである。

【図2】図2は、解析装置の構成図である。

【図3】図3は、原時系列データから周期性の強さを抽出する処理の説明図である。

【図4】図4は、時系列データを変換時系列データに変換する処理の説明図である。

【図5】図5は、時系列データの例を示す図である。

【図6】図6Aは、Fourier-CCMの解析結果を示す図である。図6Bは、Normal-CCMの解析結果を示す図である。

10

【図7】図7Aは、Fourier-CCMの解析結果を示す図である。図7Bは、Normal-CCMの解析結果を示す図である。

【図8】図8Aは、Fourier-CCMの解析結果を示す図である。図8Bは、Normal-CCMの解析結果を示す図である。

【図9】図9は、解析装置による処理手順を示すフローチャートである。

【発明を実施するための形態】

【0009】

< 1. 解析装置、解析方法、及びコンピュータプログラムの概要 >

【0010】

20

(1)実施形態にかかる解析装置は、第1解析処理を実行する。第1解析処理は、第1原時系列データから、前記第1原時系列データにおける周期性の強さの第1時系列データを生成することを含む。周期性の強さは、例えば、原時系列データにおける所定の周期の成分の強さ(振幅の大きさ)である。所定の周期は、例えば、1年周期又は1日周期である。所定の周期は、ユーザによって指定されてもよいし、解析装置によって自動的に設定されてもよい。

【0011】

第1解析処理は、第2原時系列データから、前記第2原時系列データにおける周期性の強さの第2時系列データを生成することを含む。ここで、第1時系列データの生成と第2時系列データの生成の実行順序は限定されない。すなわち、第2時系列データが第1時系列データよりも先に生成されてもよい。

30

【0012】

第1解析処理は、前記第1時系列データ及び前記第2時系列データに基づいて、因果関係解析をすることを含む。因果関係解析は、時系列データ間における因果の存在を検出できれば足りる。ここで、因果とは、第1時系列データ及び第2時系列データの一方の挙動が他方の挙動に影響することである。すなわち、因果とは、一方が原因となって、他方がその結果となる関係である。なお、両データそれぞれが他方の原因及び結果になってもよい。

【0013】

因果関係解析の手法としては、例えば、CCM(非特許文献1参照)、TE(非特許文献2参照)、状態空間モデルを用いた時系列解析(非特許文献3参照)を採用することができる。なお、因果関係解析は、時系列データ同士の因果関係の強さ又は方向を検出することを含んでもよい

40

【0014】

ここで、第1時系列データ及び第2時系列データは、いずれも周期性の強さの時系列データである。実施形態においては、第1原時系列データと第2原時系列データとの間に因果があれば、周期性の強さの第1時系列データと第2時系列データとの間にも因果が存在することを利用する。第1原時系列データと第2原時系列データとの間に因果があれば、第1時系列データ及び第2時系列データに基づいて、因果関係解析をしても、因果を検出できる。

50

## 【 0 0 1 5 】

さらに、実施形態においては、第1原時系列データと第2原時系列データとの間に因果がなく、単に周期が同調しているにすぎない場合には、原因となる時系列データにおける周期性の強さが、結果となる時系列データの周期性の強さには反映されないことを利用する。すなわち、第1原時系列データと第2原時系列データとの間に因果があれば、原因となる時系列データにおける周期性の強さが、結果となる時系列データの周期性の強さに反映される。

## 【 0 0 1 6 】

例えば、1日周期を持つ時系列データにおいて、1日周期の成分の強さは、常に一定ではなく、経時的に変化する。時系列データ間に因果があれば、原因となる時系列データにおける1日周期の成分の強さの経時変化は、結果となる時系列データにおける1日周期の成分の強さの経時変化にも反映される。一方、時系列データ間に因果がなければ、周期が同調していたとしても、1日周期の成分の強さの経時変化は、時系列データ間で異なったものとなる。

10

## 【 0 0 1 7 】

したがって、周期性の強さの第1時系列データと第2時系列データに基づいて、因果関係解析をすることで、因果がなく単に周期が同調しているにすぎない場合において、因果を誤検出することを抑制できる。

## 【 0 0 1 8 】

ここで、因果関係解析は、「前記第1時系列データ及び前記第2時系列データに基づいて」なされれば足りる。ここで、「前記第1時系列データ及び前記第2時系列データに基づいて」とは、因果関係解析を、前記第1時系列データ及び前記第2時系列データに対して行う場合のほか、派生データに対して行う場合も含む。派生データは、例えば、前記第1時系列データ及び前記第2時系列データを変換して得られたデータなど、前記第1時系列データ及び前記第2時系列データから派生して生成されたデータである。

20

## 【 0 0 1 9 】

(2) 前記第1解析処理は、前記第1時系列データを、前記第1時系列データにおける周期性の強さの変化量の時系列データである第1変換時系列データに変換することを更に含むことができる。前記第1解析処理は、前記第2時系列データを、前記第2時系列データにおける周期性の強さの変化量の時系列データである第2変換時系列データに変換することを更に含むことができる。

30

## 【 0 0 2 0 】

周期性の強さの変化量への変換により、時系列データの定常性を確保することができる。定常性の確保は、因果関係解析がCCMによる解析である場合に特に有用である。周期性の強さの変化量は、一回差分として求められてもよいし、時系列データの偏微分値として求められてもよいし、時系列データの移動平均に対する残差として求められてもよい。

## 【 0 0 2 1 】

なお、第1変換時系列データと第2変換時系列データへの変換の実行順序は限定されない。すなわち、第2変換時系列データが第1変換時系列データよりも先に生成されてもよい。

40

## 【 0 0 2 2 】

(3) 実施形態において、前記第1時系列データは、前記第1原時系列データに対して移動窓を用いたスペクトル解析をすることで生成される。また、前記第2時系列データは、前記第2原時系列データに対して移動窓を用いたスペクトル解析をすることで生成される。スペクトル解析は、例えば、フーリエ解析である。なお、移動窓の時間長さは、原時系列データが持つ周期性における周期よりも長く設定される。

## 【 0 0 2 3 】

(4) 解析装置は、前記第1原時系列データ及び前記第2原時系列データそれぞれにおける周期性の存在を識別する処理を更に実行することができる。周期性の存在を識別する処理は、前記第1原時系列データ及び前記第2原時系列データそれぞれにおける周期の長さ

50

を識別することを更に含むことができる。

【0024】

(5) 前記識別する処理の結果は、前記第1解析処理の実行の要否の判断に用いられるのが好ましい。

【0025】

(6) 前記第1解析処理は、前記第1原時系列データの周期性と前記第2原時系列データの周期性とが同調している場合に実行されるのが好ましい。

【0026】

(7) 前記第1原時系列データ及び前記第2原時系列データに対して前記因果関係解析をする第2解析処理を更に実行するのが好ましい。

10

【0027】

(8) 解析装置は、前記第1解析処理の結果及び前記第2解析処理の結果を出力するのが好ましい。

【0028】

(9) 実施形態に係る解析方法は、第1原時系列データから、前記第1原時系列データにおける周期性の強さの第1時系列データを生成し、第2原時系列データから、前記第2原時系列データにおける周期性の強さの第2時系列データを生成し、前記第1時系列データ及び前記第2時系列データに基づいて、因果関係解析をよい。コンピュータは、プロセッサ300が、記憶装置200に記憶されたコンピュータプログラム251を実行することで、解析装置100として機能する。実施形態のコンピュータプログラム251は、図1及び図9に示す処理をコンピュータに実行させるよう構成されている。

20

【0029】

プロセッサ300は、コンピュータプログラム251を実行することにより、図2に示す様々な処理310、320、330、340、350、360を実行することができる。これらの処理310、320、330、340、350、360については後述する。

【0030】

解析装置100には、第1原時系列データ211と第2原時系列データ212とが外部から入力され、記憶装置200に保存される。第1原時系列データ211と第2原時系列データ212とは、因果関係の検出対象となる時系列データである。

【0031】

図1に示すように、第1解析処理310では、プロセッサ300が、第1原時系列データ211と第2原時系列データ212とを記憶装置200から読み出す(ステップS11)。プロセッサ300は、第1原時系列データ211と第2原時系列データ212とから、第1時系列データ221と第2時系列データ222とを生成する生成処理を実行する(ステップS12)。

30

【0032】

生成処理では、図3に例示する第1原時系列データ211から、図4に例示する第1時系列データ221が生成される。同様に、第2原時系列データ212から、第2時系列データ222が生成される。第1時系列データ221及び第2時系列データ222は、それぞれ、周期性の強さの時系列データである。ここで、図3に示す第1原時系列データ211において、横軸は時間(年)であり、縦軸は観測値である。図3に示す第1時系列データ221は、2週間に一度観測された10年分のデータであり、1年の周期性を持つ。

40

【0033】

実施形態において、周期性の強さの第1時系列データ221及び第2時系列データ222は、移動窓(moving window)Wを用いたフーリエ解析により生成される。図3において、移動窓の時間長さは、3年に設定されている。移動窓の時間長さは、因果関係解析の対象となる時系列データの全時間長さよりも短く、その時系列データの持つ周期性における周期よりも長く設定される。

【0034】

図3は、移動3年窓(Moving 3-year window)Wを適用したフーリエ解析結果214も

50

示している。フーリエ解析結果 2 1 4 において、横軸は周波数であり、縦軸はパワー（振幅）である。図 3 において、周波数が 1 である位置が、1 年周期(1-year periodicity)に対応している。生成処理では、フーリエ解析結果 2 1 4 から、1 年周期のパワー(Power of 1-year periodicity)が抽出される。1 年周期のパワーは、第 1 原時系列データ 2 1 1 における 1 年周期の強さを示している。以下では、1 年周期のパワーを、季節性の強さ A ともいう。

【 0 0 3 5 】

移動窓 W の適用位置を時間軸に沿った移動方向 m へ徐々にずらすことで、図 4 に示すように、7 年分の「1 年周期のパワー（強さ）」の時系列データである第 1 時系列データ 2 2 1 が得られる。第 2 時系列データ 2 2 2 も同様にして得られる。生成された第 1 時系列データ 2 2 1 及び第 2 時系列データ 2 2 2 は、記憶装置 2 0 0 に保存される。

10

【 0 0 3 6 】

続いて、プロセッサ 3 0 0 は、第 1 時系列データ 2 2 1 と第 2 原時系列データ 2 2 2 とから、第 1 変換時系列データ 2 3 1 と第 2 変換時系列データ 2 3 2 とを生成する変換処理を実行する（ステップ S 1 3）。変換処理は、記憶装置 2 0 0 に保存されているデータ 2 2 1, 2 2 2 を読みだして行われる。第 1 変換時系列データ 2 3 1 及び第 2 変換時系列データ 2 3 2 は、それぞれ、周期性の強さの変化量の時系列データである。変換処理は、時系列データの定常性を確保するために行われる。変換処理は行われなくてもよい。

【 0 0 3 7 】

実施形態において、第 1 変換時系列データ 2 3 1 及び第 2 変換時系列データ 2 3 2 は、時系列データ 2 2 1, 2 2 2 において、例えば、ある時間 t における季節性の強さ A[t] と、時間的に直前の時点 t-1 における季節性の強さ A[t-1] との 1 回差分をとることで生成される。すなわち、1 回差分は、 $A[t]-A[t-1]$  によって求められる。1 回差分をデータ 2 2 1 の全時間範囲についてとることで、図 4 に示す第 1 変換時系列データ 2 3 1 が得られる。第 2 変換時系列データ 2 3 2 についても同様にして得られる。得られた第 1 変換時系列データ 2 3 1 及び第 2 変換時系列データ 2 3 2 は、記憶装置 2 0 0 に保存される。

20

【 0 0 3 8 】

プロセッサ 3 0 0 は、第 1 変換時系列データ 2 3 1 及び第 2 変換時系列データ 2 3 2 に対して、CCM による解析を実行する。なお、ステップ S 1 3 の変換処理を省略した場合、プロセッサ 3 0 0 は、第 1 時系列データ 2 2 1 と第 2 時系列データ 2 2 2 に対して、CCM による解析を実行する。CCM による解析は、記憶装置 2 0 0 に保存されているデータ 2 3 1, 2 3 2 又はデータ 2 2 1, 2 2 2 を読みだして行われる。

30

【 0 0 3 9 】

< 2 . 2 実験 >

【 0 0 4 0 】

第 1 解析処理(Fourier-CCM) 3 0 0 の有効性を確認する実験をシミュレーションにより行った。実験のため時系列データ 2 2 2 を人工的に生成した。図 5 に示すように、実験では、第 1 原時系列データ 2 1 1 として、気象要因 A (Climate variable A) の時系列データ、気象要因 B (Climate variable B) の時系列データ、気象要因 C (Climate variable C) の時系列データを作成した。また、第 2 原時系列データ 2 1 2 として、被説明変数 Y (Explained variable Y) の時系列データを作成した。なお、図 5 には、被説明変数 Y の内部状態 X を示す時系列データも示した。

40

【 0 0 4 1 】

気象要因 A、気象要因 B、気象要因 C、及び被説明変数 Y それぞれの時系列データは、1 年周期を持つ。なお、図 5 では、1 年周期の正弦波 S が、気象要因 A、内部状態 X、及び被説明変数 Y の時系列データに重ねて描かれている。

【 0 0 4 2 】

図 5 に示すように、気象要因 A は、コンスタントに被説明変数 Y に影響を与えている。すなわち、気象要因 A は、被説明変数 Y の季節性を生み出す原因になっている。気象要因 B は、被説明変数 Y に影響を与えているものの、その効果は、特定の時季に限定されてお

50

り、被説明変数 Y における季節性全体を作り出しているわけではない。気象要因 C は、気象要因 A , B と同じく 1 年周期を持つが、被説明変数 Y には影響していない。すなわち、気象要因 C は、被説明変数 Y に同調しているが、被説明変数 Y との間に因果があるわけではない。

【 0 0 4 3 】

気象要因 A , B , C の時系列データの生成の仕方は、以下のとおりである。

- 1 . 生成するデータの長さ (= 観測期間) を設定する ( 月 2 回観測の 10 年分であれば、 $24 \times 10 = 240$  データ点 ) 。
- 2 . 正規分布からランダムに値をデータ点分 ( 240 個 ) 生成する。
- 3 . ランダムな値の Exponential をとり、気象の「もと時系列」とする。
- 4 . 「もと時系列」に季節性 (= 1 年周期の正弦波 S ) を足し、周期性 ( 1 年周期 ) をもつ時系列とする。ここで、季節性の強さは正弦波 S の値とランダムな値の相対的な重要性によって変化させることができる。
- 5 . 「季節性の強さ」に多少の変動を加えるために、正弦波 S の振幅を時期によって少しずつ変動させる。例えば、気象要因 A の時間点 40 付近の振幅は、時間点 200 付近の振幅よりもやや大きい。「季節性の強さの変動」は気象要因 A , B , C それぞれでバラバラとした。ただし、正弦波の周期は、気象要因 A , B , C において同じである。

10

【 0 0 4 4 】

被説明変数 Y の時系列データの生成の仕方は、以下のとおりである。

- 1 . 生成するデータの長さ (= 観測期間) を設定する ( 月 2 回観測の 10 年分なら、 $24 \times 10 = 240$  データ点 ) 。
- 2 . ロジスティック写像 ( カオス動態 ) の微分方程式を利用して、決定論的な動態 ( 内部状態 X ) を生成する ( $X[t+1] = X[t] * (r + s * X[t])$  を利用 ) 。
- 3 . 気象要因 A から継続的で一定の影響を受け、気象要因 B からは 1 年の内ある時季のみ影響を受け ( したがって、振幅全体の決定には関わらない )、気象要因 C からは影響をうけないと仮定して、被説明変数 Y の時系列を生成する。

20

【 0 0 4 5 】

以上のような周期性を有する A , B , C 及び Y の時系列データに基づいて、A , B , C のうちの気象要因が、Y の季節性を生み出す主たる要因かを特定できれば、第 1 解析処理 3 0 0 の有効性が示される。

30

【 0 0 4 6 】

図 6 A , 図 6 B , 図 7 A , 図 7 B , 図 8 A , 図 8 B は、第 1 解析処理 (Fourier-CCM) 3 0 0 及び通常の CCM (Normal-CCM; 通常の CCM は、特許文献 1 に示す CCM と同様) の実験結果を示している。第 1 解析処理 (Fourier-CCM) 3 0 0 及び通常の CCM の実験条件は、以下のとおりである。

【 0 0 4 7 】

< 季節性 (Seasonality) の強さ >

気象要因 A , B , C 及び被説明変数 Y の各時系列データにおいて、季節性の強さを 0 , 1 , 2 に変化させた。季節性の強さは、季節性を示す正弦波 S の振幅の大きさとランダムな値の標準偏差との比を示している。季節性の強さが 0 であれば、正弦波 S の振幅は 0 であり、季節性がないことを示す。季節性の強さが 1 であれば、ランダムな値の標準偏差と同じ振幅の正弦波 S によって季節性が生じている。季節性の強さが 2 であれば、ランダムな値の標準偏差の 2 倍の振幅の正弦波 S によって季節性が生じており、季節性が非常に強いことを示す。

40

【 0 0 4 8 】

< 観測誤差 (Observation error) >

気象要因 A , B , C 及び被説明変数 Y の各時系列データに、観測誤差 ( 正規分布を仮定 ) を足した。観測誤差の強さは、標準偏差の 0 , 5 , 1 0 , 1 5 , 2 0 , 2 5 , 3 0 , 3 5 , 4 0 , 4 5 , 5 0 % に変化させた。

【 0 0 4 9 】

50



< 実験の繰り返し条件 >

第1解析処理(Fourier-CCM)300は、季節性の強さと観測誤差のパラメータセット(一つの季節性の強さと一つの観測誤差のセット)について繰り返しテストを行った。1回のテストあたり、1000個のサロゲートデータ(シャッフル時系列データ)を生成し、それぞれに第1解析処理(Fourier-CCM)300及び通常のCCMをした。

【0050】

図6A, 図6B, 図7A, 図7B, 図8A及び図8Bにおいて、横軸は観測誤差であり、縦軸は、combined p-valueを示している。combined p-valueは、いわゆるp値と同様の概念であり、例えば、0.10未満、より好ましくは0.05未満であれば、説明変数である気象要因が、被説明変数Yの予測に有効であること、すなわち、気象要因と被説明変数Yとに因果があることが検出されていることになる。

10

【0051】

図6Aは、気象要因Aの時系列データを第1原時系列データとし、被説明変数Yの時系列データを第2原時系列データとして、第1解析処理(Fourier-CCM)300を行った結果を示す。図6Bは、気象要因Aの時系列データと被説明変数Yの時系列データに対して、通常のCCM(Normal-CCM; 通常のCCMは、特許文献1に示すCCMと同様)を行った結果を示す。

【0052】

コンスタントに被説明変数Yに影響を与えている気象要因Aについては、図6Aに示すように、気象要因Aに季節性がある(Seasonality=1又は2)場合、観測誤差が大きすぎない範囲(例えば、観測誤差が0.3未満)において、気象要因Aと被説明変数Yとの因果が第1解析処理(Fourier-CCM)300によって検出されている。なお、気象要因Aに季節性がない(Seasonality=0)場合、第1解析処理(Fourier-CCM)300では、季節要因Aと被説明変数Yとの因果は検出されない。

20

【0053】

一方、通常のCCMによれば、図6Bに示すように、コンスタントに被説明変数Yに影響を与えている気象要因Aについては、季節性の有無にかかわらず、気象要因Aと被説明変数Yとの因果を検出することができる。

【0054】

図7Aは、気象要因Bの時系列データを第1原時系列データとし、被説明変数Yの時系列データを第2原時系列データとして、第1解析処理(Fourier-CCM)300を行った結果を示す。図7Bは、気象要因Bの時系列データと被説明変数Yの時系列データに対して、通常のCCMを行った結果を示す。

30

【0055】

被説明変数Yに影響を与えているものの、その効果が、特定の時季に限定されている気象要因Bについては、通常のCCMによると、気象要因Bと被説明変数Yとの因果が誤検出されることがある。すなわち、気象要因Bの季節性が被説明変数Yと同調しているため、気象要因Bの季節性が強い(Seasonality=2)場合、観測誤差が十分に小さい範囲(例えば、観測誤差が0.1未満)において、気象要因Bと被説明変数Yとの因果が誤検出されている。これに対して、第1解析処理(Fourier-CCM)300によれば、季節性の強さにかかわらず、気象要因Bと被説明変数Yとの因果は検出されておらず、通常のCCMのような誤検出が防止されていることがわかる。

40

【0056】

図7Cは、気象要因Cの時系列データを第1原時系列データとし、被説明変数Yの時系列データを第2原時系列データとして、第1解析処理(Fourier-CCM)300を行った結果を示す。図7Cは、気象要因Cの時系列データと被説明変数Yの時系列データに対して、通常のCCMを行った結果を示す。

【0057】

被説明変数Yと同じく1年周期を持つが、被説明変数Yには影響していない気象要因Cについても、気象要因Cの季節性が強くなるほど、Combined p-valueの値が低下しており、気象要因Cと被説明変数Yとの因果の誤検出の可能性が示唆されている。これに対して

50

、第1解析処理(Fourier-CCM)300によれば、季節性の強さでCombined p-valueの値がほとんど変化しておらず、周期の同調によって因果を誤検出してしまう可能性が低いことがわかる。

【0058】

以上のように、通常のCCMであると、周期の同調性のため、気象要因A, B, Cのいずれも被説明変数Yとの因果が検出されてしまうことがあるのに対して、第1解析処理(Fourier-CCM)300によれば、被説明変数Yとの間に因果があるのは気象要因Aであると正しく検出することができる。すなわち、気象要因A, B, Cのうち気象要因Aが、被説明変数Yの季節性を生み出す主たる要因かを特定することができる。

【0059】

なお、気象要因に季節性がない場合、図6Aのように、第1解析処理(Fourier-CCM)300では、因果を検出できないのに対して、図6Bに示すように、通常のCCMでは、因果を検出できる。したがって、第1解析処理(Fourier-CCM)300及び通常のCCMを併用するのも有効であることがわかる。

【0060】

< 2.3 解析装置における他の処理 >

【0061】

図9は、解析装置10において、第1解析処理に関連して行われる処理を示している。ステップS21において、プロセッサ300は、第1原時系列データ211に対する周期性識別処理330を実行する。周期性識別処理330は、第1原時系列データ211全体をフーリエ解析することを含む。フーリエ解析により、第1原時系列データ211において、パワーが所定の閾値よりも大きくなる周期が存在すれば、プロセッサ300は、第1原時系列データ211では、その周期での周期性が存在することを識別できる。また、プロセッサ300は、フーリエ解析により、その周期の長さを識別することもできる。

【0062】

ステップS22において、プロセッサ300は、第2原時系列データ212に対する周期性識別処理330を実行する。第2原時系列データ212に対する周期性識別処理330は、第1原時系列データ211に対する周期性識別処理330と同様である。周期性識別処理330によって、第2原時系列データ212における周期性の存在の識別と周期の長さの識別とが可能である。

【0063】

ステップS23において、プロセッサ300は、解析手法判断処理340を実行する。実施形態においては、解析手法判断処理340では、第1原時系列データ211の周期性と第2原時系列データ212の周期性とが同調しているか否かが判断される。ステップS23においては、第1原時系列データ211の周期の長さとは第2原時系列データ212の周期の長さを比較し、一致していれば、同調していると判断される。なお、周期の長さの比較においては、長さが概ね一致していれば、一致と判断してよい。

【0064】

ステップS23において同調していると判断された場合、プロセッサ300は、ステップS24において、第1解析処理(Fourier-CCM)310を実行する。第1解析処理(Fourier-CCM)310は、時系列データ同士が同調している場合の解析手法として有効である。解析手法判断処理340においては、識別処理330の結果を用いて、第1解析処理310の実行の要否が判断されるため、必要に応じて第1解析処理310を実行することができる。

【0065】

一方、ステップS23において非同調と判断された場合、プロセッサ300は、ステップS25において、第2解析処理(CCM)320を実行する。第2解析処理(CCM)320は、第1原時系列データ211及び第2原時系列データ212に対して通常のCCMによる解析をする。時系列データ同士が同調していない場合には、通常のCCMによる解析が有効である(図6A及び図6参照)。なお、ステップS24においても、第1解析処理として

10

20

30

40

50

、Fourier-CCMに加えて、通常のCCMによる解析を実行してもよい。

【0066】

ステップS26において、プロセッサ300は、第1解析処理310又は第2解析処理の結果の出力処理350を実行する。解析処理の結果は、因果関係の存在の有無を含むことができる。解析処理の結果は、因果関係の強さ及び方法を含んでもよい。

【0067】

因果関係の解析として、Fourier-CCMとCCMの双方が実行される場合、出力処理においては、双方の解析結果を出力してもよい。双方の解析結果が出力されることで、ユーザは双方の解析結果を比較することができる。

【0068】

第1解析処理310又は第2解析処理320により確認された因果関係は、記憶装置200において、因果関係データ241として保存される。因果関係データ241を用いることで、因果関係における原因となる要因に変動が生じた場合に、その原因が影響する要因における変動を予測する予測処理360を行うことができる。

【0069】

< 2.4 応用 >

【0070】

実施形態に係る解析処理310、320により検出された因果関係データを用いた予測は、例えば、次のようなものである。

- ・腸内細菌に関する時系列データから、人間の健康状態の予測
- ・水中の微生物に関する時系列データから、魚の病気・藻類の発生の予測
- ・気象要因、土中/水中の栄養素、及び微生物の量それぞれの時系列データから、植物の生育の予測
- ・気象要因の時系列データから、森林の植生・生理状態(CO<sub>2</sub>固定能力の変化など)の予測
- ・物理現象(振動・温度)の時系列データから、機械の故障の予測
- ・物理現象(振動・温度)の時系列データから、製品の品質の予測
- ・生体現象(血圧・体温)の時系列データから、病気の予測
- ・気象要因の時系列データから、感染症の発生の予測(例えば、空気の乾燥からインフルエンザ流行などを予測)
- ・気象要因の時系列データから、害虫(農業害虫・蚊)の発生の予測
- ・環境要因の時系列データから、人間の概日リズムへの影響の予測
- ・気象要因の時系列データから、人間の購買行動への影響の予測(例えば、気温から夏にビールが良く売れることなどを予測)
- ・遺伝子発現量の変動の時系列データから、生物の生理状態(病気など)の予測
- ・遺伝子同士の相互作用ネットワークの推定

【0071】

< 3. 付記 >

本発明は、上記実施形態に限定されるものではなく、様々な変形が可能である。

【符号の説明】

【0072】

- 100 解析装置
- 200 記憶装置
- 211 第1原時系列データ
- 212 第2原時系列データ
- 214 フーリエ解析結果
- 221 第1時系列データ
- 222 第2時系列データ
- 231 第1変換時系列データ
- 241 第2変換時系列データ

10

20

30

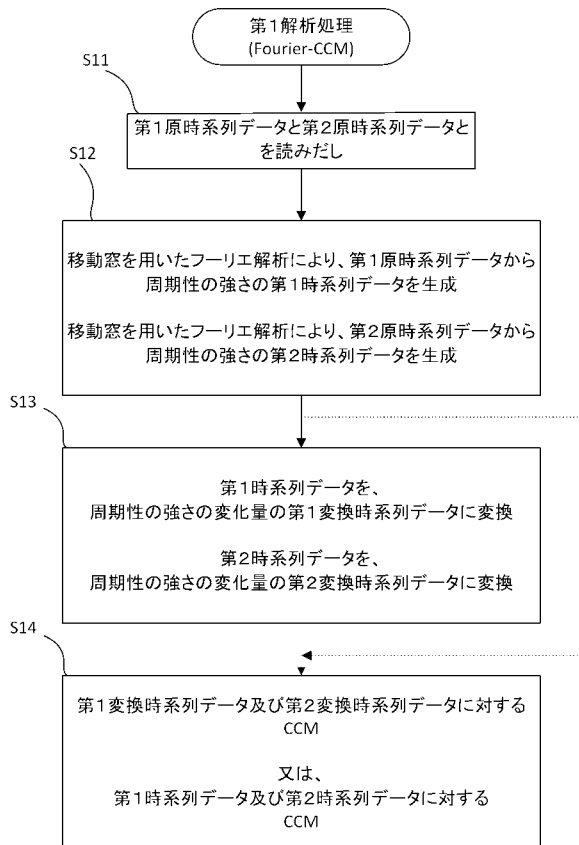
40

50

- 2 5 1 コンピュータプログラム
- 3 0 0 プロセッサ
- 3 1 0 第1解析処理 (Fourier-CCM)
- 3 2 0 第2解析処理 (CCM)
- 3 3 0 周期性識別処理
- 3 4 0 解析手法判断処理
- 3 5 0 出力処理
- 3 6 0 予測処理
- S 正弦波
- W 移動窓
- m 移動方向
- A 季節性の強さ

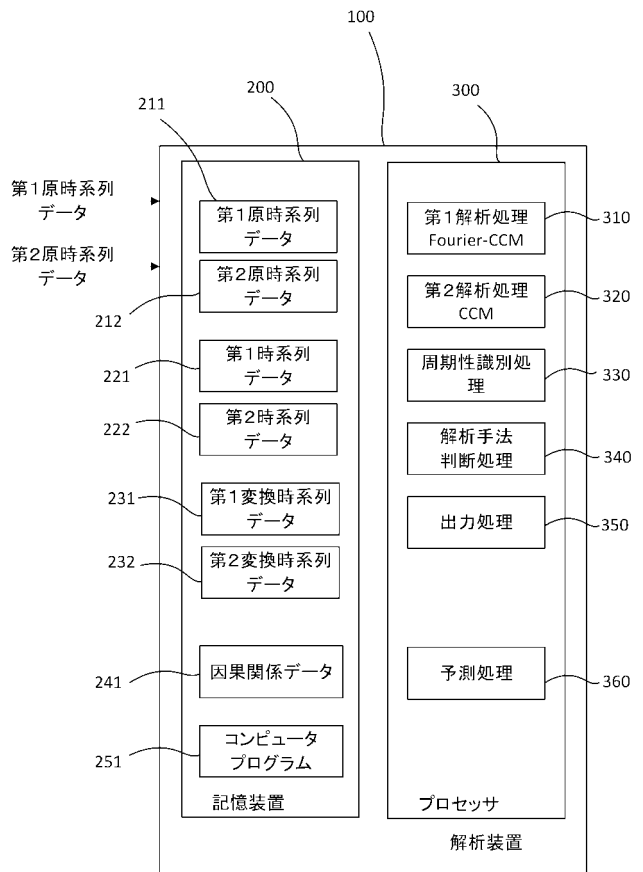
【 図 1 】

図1

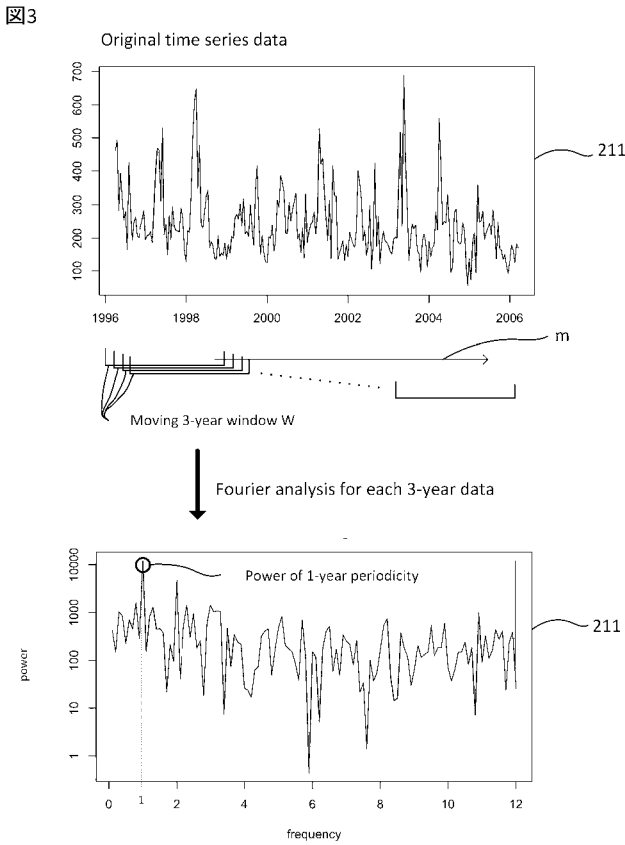


【 図 2 】

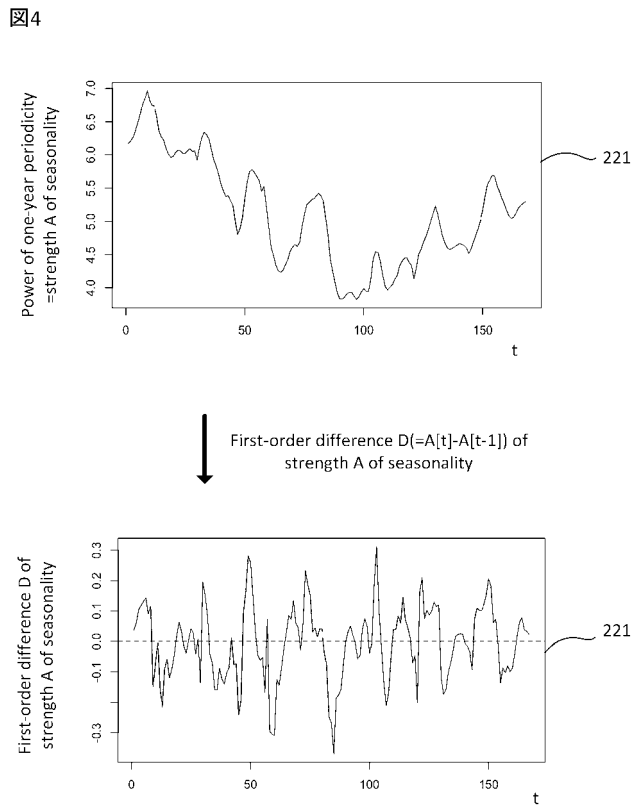
図2



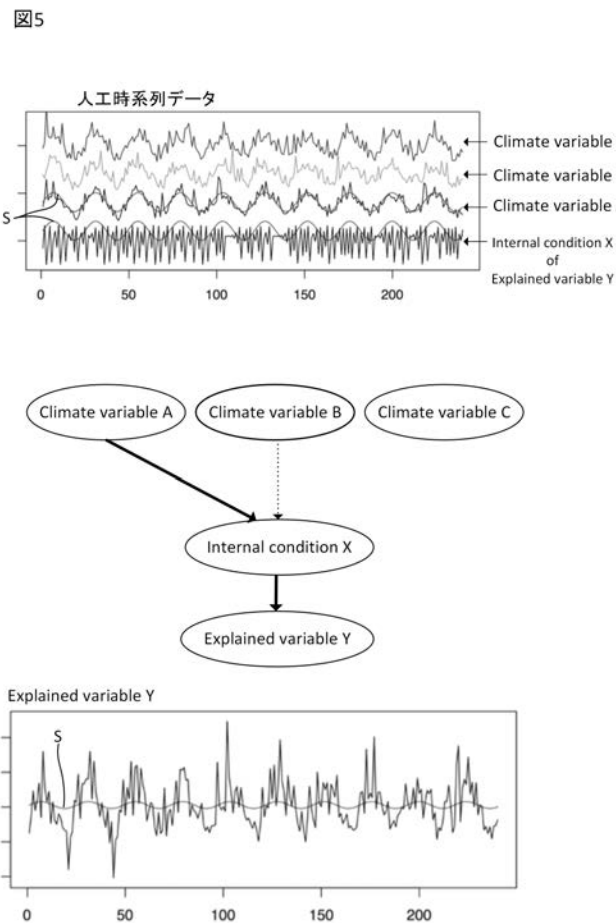
【 図 3 】



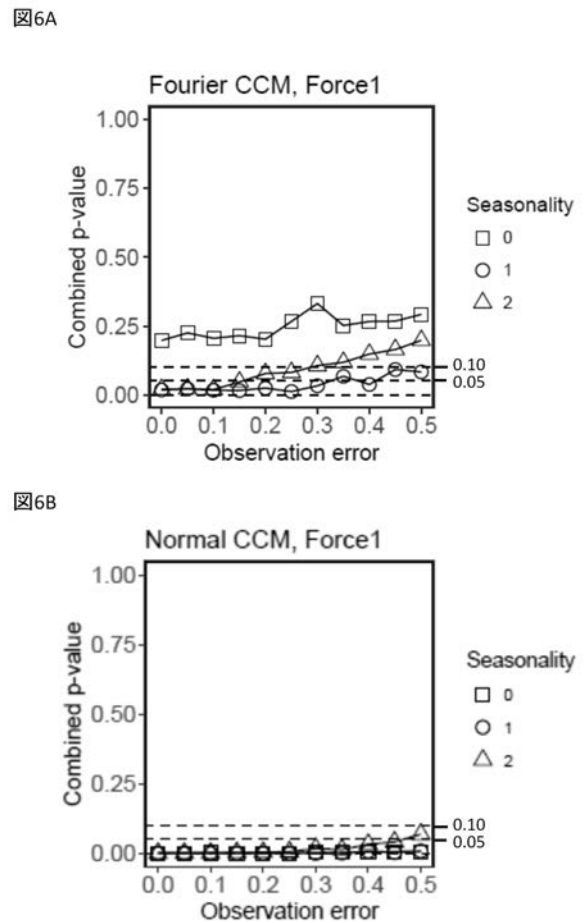
【 図 4 】



【 図 5 】



【 図 6 】



【 図 7 】

図7A

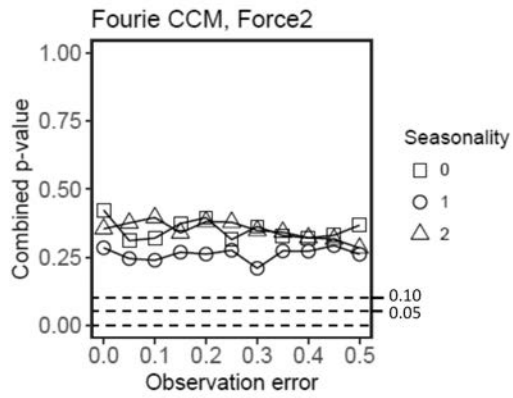
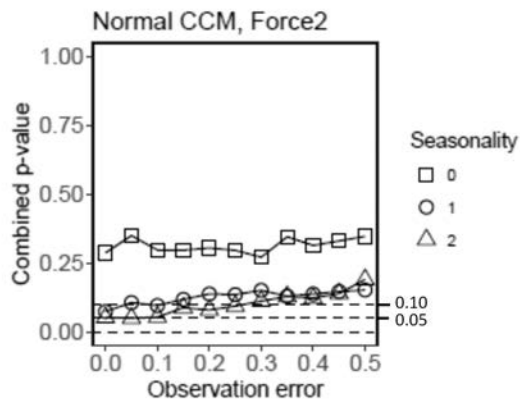


図7B



【 図 8 】

図8A

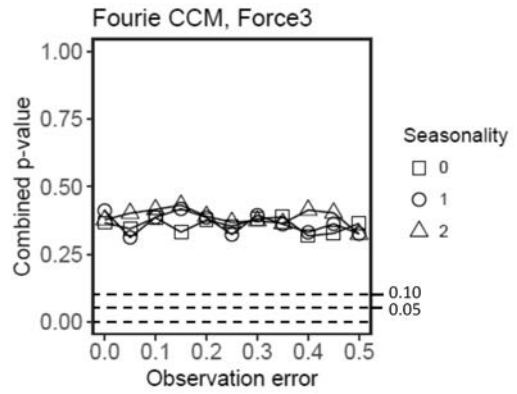
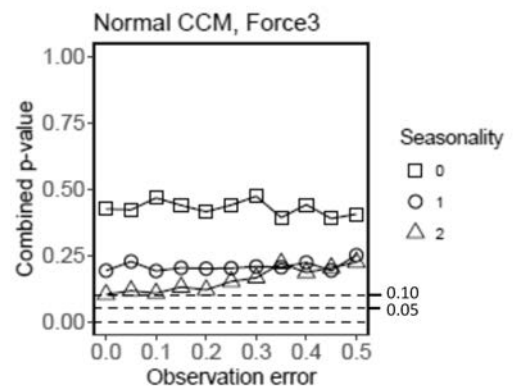


図8B



【 図 9 】

図9

