International Journal of Data Warehousing and Mining
Volume 18 • Issue 2

Concept of Temporal Pretopology for the Analysis for Structural Changes:
Application to Econometrics

Nazha Selmaoui-Folcher, University of New Caledonia, New Caledonia*
https://orcid.org/0000-0003-1667-3819
Jannai Tokotoko, University of New Caledonia, New Caledonia
Samuel Gorohouna, University of New Caledonia, New Caledonia
Laisa Roi, University of New Caledonia, New Caledonia
Claire Leschi, INSA de Lyon, France
Catherine Ris, University of New Caledonia, New Caledonia

ABSTRACT

Pretopology is a mathematical model developed from a weakening of the topological axiomatic. It was initially used in economic, social, and biological sciences and next in pattern recognition and image analysis. More recently, it has been applied to the analysis of complex networks. Pretopology enables to work in a mathematical framework with weak properties, and its nonidempotent operator called pseudo-closure permits to implement iterative algorithms. It proposes a formalism that generalizes graph theory concepts and allows to model problems universally. In this paper, the authors will extend this mathematical model to analyze complex data with spatiotemporal dimensions. The authors define the notion of a temporal pretopology based on a temporal function. They give an example of temporal function based on a binary relation and construct a temporal pretopology. They define two new notions of temporal substructures which aim at representing evolution of substructures. They propose algorithms to extract these substructures. They experiment the proposition on two data and two economic real data.

KEYWORDS

Modelisation, Pretopological Model, Propagation Dynamics, Pseudo-Closure, Sectoral Econometrics, Structural Analysis, Structural Evolution, Temporal, Temporal Data

INTRODUCTION

Structural analysis of complex networks enables to explore social networks, sector economics, lexical taxonomies, etc. Networks are generally represented by graphs, and the aim is to study relevant interactions between individuals (nodes). Those interactions can be of various types such as dependency, influence, etc.

However, the graph theory framework does not allow to easily model relations at different levels (e.g. individual to group, group to group), thus limiting both modelling possibilities and levels at

DOI: 10.4018/IJDWM.298004 *Corresponding Author

which structural analysis can be performed (Dalud-Vincent et al., 2001). To overcome such limitations, (Dalud-Vincent et al., 2001) proposed to use the pretopological framework, presented in (Belmandt, 1993; Brissaud et al., 2011) as an extension of graph theory.

The concept of pretopology was first introduced in the 1970's (Brissaud, 1971, 1975). It results from a weakening of the axiomatic of topology, the latter being seen as a mathematical way to formalize human perception based on the concept of similarity. The notion of similarity can be declined mainly in two ways: proximity and approximation. Proximity can be a relation between two objects, an object and a set of objects, or between two sets of objects. On the other hand, given a discrete set E, the notion of closure ad applied to a point $x \in E$ (resp. a set of points $A\ E$) enables a simple formalization of the concept of approximation: $ad\left(x\right)$ (resp. $ad\left(A\right)$) is what one sees when looking at x (resp. A). Doing so, it provides a means to implement how humans are perceiving patterns. However, in both cases, constraints brought by the axiomatic of topology often fail to process real data efficiently. In particular, the idempotence of the operator *ad* allows to generate only one possible approximation of an individual or a set. Unlike topology, a pretopology is defined by a function called pseudo-closure which is not (necessarily) idempotent. It thus offers the opportunity to follow an approximation process step by step as well as to model the notion of perception threshold.

Further developed from the 1980's (Auray et al., 1979; Duru, 1980; Hubert Emptoz, 1983), the pretopological framework permits to study weak topological structures, in particular discrete and finite structures based on models generated step by step (propagation phenomenon), and describing for example information spreading in complex networks.

Pretopology found its first applications in social sciences and econometrics (Auray et al., 1979; Duru, 1980), social networks (Basileu et al., 2012; Bui, 2018; Dalud-Vincent et al., 2001; Levorato, 2011), pattern recognition (Hubert Emptoz, 1983), and image analysis (Arnaud et al., 1986; Lamure, 1987; Piegay, 1997; Piegay et al., 1995; Selmaoui et al., 1993). More recently, researchers have brought the pretopological framework in domains such as machine learning (Le et al., 2007) or text exploration (Cleuziou et al., 2011).

More generally, pretopology has proved to be of great interest for building mathematical models adapted to data set structures in order to carry out data structural analysis, to extract tendencies (clustering), or to predict events (supervised classification). In those contexts, pretopology is mainly used to process static data. Our present contribution aims to extend the pretopological framework to the domain of spatio-temporal data analysis. We first establish a formalism for a temporal pretopology in order to build a theoretical foundation for both temporal and structural evolutions. Next, we provide some examples of evolution based on a temporal pseudo-closure function. Finally, we propose a first application of our formalism to the evolution of sectoral economics data by modelling the underlying dynamics through a sequence of pretopological spaces.

BASIC DEFINITIONS AND FORMALISM

Our contribution is aimed at introducing a temporal dimension in a pretopological space. Before presenting our approach, we briefly recall definitions and basic concepts of a pretopological model. For further details, the reader can refer to (Belmandt, 1993; Brissaud et al., 2011; Bui, 2018).

Let E be a non-empty set of individuals or objects. Let $\mathcal{P}\left(E\right)$ be the set of subsets of E. Let us define on E an extension operator *ad* from $\mathcal{P}\left(E\right)$ to $\mathcal{P}\left(E\right)$ associated to a dual operator *int*.

Definition 1 (Pretopological space and pseudo-closure) $\left(E,\boldsymbol{ad}\right)$ is a pretopological space if and only if $\boldsymbol{ad}$ is an operator from $\mathcal{P}\left(E\right)$ to $\mathcal{P}\left(E\right)$ called **pseudo-closure** and verifying:

- $\boldsymbol{ad}\left(\varnothing\right) = \varnothing$

- $\forall\ A\in \mathcal{P}\left(E\right),\quad A\subset ad\left(A\right)$

This definition can be given by the dual operator ***int*** called **pseudo-opening** and verifying:

- $\boldsymbol{int}\left(\varnothing\right)=\varnothing$
- $\forall\ \boldsymbol{A}\in \boldsymbol{\mathcal{P}}\left(\boldsymbol{E}\right),\quad \boldsymbol{int}\left(\boldsymbol{A}\right)\subset \boldsymbol{A}$

The dual operator ***int*** verifies $\forall A\in \mathcal{P}\left(E\right),\ \boldsymbol{int}\left(A\right)=\overline{\boldsymbol{ad}\left(\bar{A}\right)}$ where $\bar{A}$ is the complement of the set A.

Remarks:

1. The pseudo-closure *ad* is not necessarily idempotent, i.e. we have: $A\subset ad\left(A\right)\subset ad\left[ad\left(A\right)\right]\subset\ldots$, whereas it is the case for a topological space. This property of non-idempotence allows to design iterative operators and consequently iterative algorithms.
2. Operators *ad* and int being dual, often only ad is actually defined.

Definition 2 Let $\left(E,\boldsymbol{ad}\right)$ be a pretopological space, let $n\in\mathbb{N}^{*}$ be a positive integer. $\forall A\subset E$, we designate by $ad^{n}\left(A\right)$ the composition of ad n times, i.e. $ad^{n}\left(A\right)=ad\underbrace{\left(ad\left(\ldots ad\left(A\right)..\right)\right)}_{n\ times}$.

Definition 3 (Closed subset) Let A be a subset of E. A is called **closed subset** if and only if $ad\left(A\right)=A$.

Definition 4 (Closure) Let A be a subset of E, the **closure** of A in $\left(E,ad\right)$ is the smallest closed subset noted $F\left(A\right)$ verifying: $A\subset F\left(A\right)$ et $\exists p\geq 1$ such that $F\left(A\right)=ad^{p}\left(A\right)=ad^{p+1}\left(A\right)$.

Definition 5 (Elementary closed subset) $\forall x\in E$, we call **elementary closed subset** associated to x, noted $F\left(x\right)$ the closure of subset $\left\{x\right\}$, which verifies: $\exists p\geq 1$, such that $F\left(x\right)=ad^{p}\left(\left\{x\right\}\right)=ad^{p+1}\left(\left\{x\right\}\right)$

Types Of Pretopological Spaces

Based on the definition of its pseudo-closure, a pretopological space may verify some interesting properties. Those properties are used to define different types of pretopological spaces (Auray et al., 1979; Belmandt, 1993; Brissaud et al., 2011). Some of them are presented below.

Definition 6 (Pretopological space of type $\mathcal{V}$) A pretopological space $\left(E,ad\right)$ is of type $\mathcal{V}$ if and only if

$$\forall A,B\in \mathcal{P}\left(E\right),\qquad \left(A\subseteq B\right)\Rightarrow\left(ad\left(A\subseteq ad\left(B\right)\right)\right.$$

Definition 7 (Pretopological space of type $\mathcal{V}_{D}$) A pretopological space $\left(E,ad\right)$ is of type $\mathcal{V}_{D}$ if and only if

$$\forall A,B\in \mathcal{P}\left(E\right),\quad ad\left(A\cup B\right)=ad\left(A\right)\cup ad\left(B\right)$$

Definition 8 (Pretopological space of type $\mathcal{V}_{DS}$) A pretopological space $\left(E,ad\right)$ is of type $\mathcal{V}_{DS}$ if and only if, for any set family $\left(A_i\right)_{i\in I}$ of E, we have

$$ad\left(U_{i\in I}A_i\right)=U_{i\in E}ad\left(A_i\right)$$

Definition 9 (Topological space) (E, ad) is a topological space if and only if (E, ad) is a pretopological space of type $\mathcal{V}_D$, verifying, in addition the idempotence axiom: $\forall A\in\mathcal{P}\left(E\right),\ ad\left(ad\left(A\right)\right)=ad\left(A\right)$

Proposition 1 Any pretopological space of type $\mathcal{V}_D$ is a pretopological space of type $\mathcal{V}$.

Example 1 (ε-neighbors pretopology) Let ε be a positive real number, *d* a distance measure (or dissimilarity etc.) defined on E. We can define a pretopological space $\left(E,ad_\varepsilon\right)$ where:

- $\boldsymbol{ad}_\varepsilon\left(\varnothing\right)=\varnothing$
- $\forall \boldsymbol{x}\in \boldsymbol{E},\ \boldsymbol{ad}_\varepsilon\left(\left\{\boldsymbol{x}\right\}\right)=\left\{\boldsymbol{y}\in \boldsymbol{E}\ /\ \boldsymbol{d}\left(\boldsymbol{x},\boldsymbol{y}\right)\le\varepsilon\right\}$
- $\forall \boldsymbol{xA}\subset \boldsymbol{E},\ \boldsymbol{ad}_\varepsilon\left(\boldsymbol{A}\right)=U_{\boldsymbol{x}\in\boldsymbol{A}}\boldsymbol{ad}_\varepsilon\left(\left\{\boldsymbol{x}\right\}\right)$

Property 1: $\left(E,ad_\varepsilon\right)$ is a pretopological space of type $\mathcal{V}_{DS}$.

This example of pre-topology was used by (Hubert Emptoz, 1983) to develop a clustering method having the same principle as the DBSCAN method (Ester et al., 1996). This method is based on the propagation of pseudo-closure from a germination point. The germination points are chosen by their high density around them. The definition of density used in this method is the structuring function also defined by (Emptoz et al., 1981).

Example 2 (k nearest neighbors (knn) pretopology) Let k be a positive integer, *d* a distance measure (or dissimilarity etc.) defined on E. We can define a pretopological space $\left(E,ad_{knn}\right)$ where:

- $\boldsymbol{ad}_{knn}\left(\varnothing\right)=\varnothing$
- $\forall \boldsymbol{x}\in \boldsymbol{E},\boldsymbol{ad}_{knn}\left(\left\{\boldsymbol{x}\right\}\right)=\left\{\boldsymbol{y}\in \boldsymbol{E}\ /\ y\ \text{ is a k nearest neighbors of x}\right\}$
- $\forall \boldsymbol{xA}\subset \boldsymbol{E},\ \boldsymbol{ad}_{knn}\left(\boldsymbol{A}\right)=U_{\boldsymbol{x}\in\boldsymbol{A}}\boldsymbol{ad}_{knn}\left(\left\{\boldsymbol{x}\right\}\right)$

Example 3 (k-adjacency pretopology) Let k be a positive integer, *d* a distance measure (or dissimilarity etc.) defined on E. We can define a pretopological space $\left(E,ad_{k-adjacency}\right)$ where:

- $\boldsymbol{ad}_{k-adjacency}\left(\varnothing\right)=\varnothing$
- $\forall \boldsymbol{x}\in \boldsymbol{E},\boldsymbol{ad}_{k-adjacency}\left(\left\{\boldsymbol{x}\right\}\right)=\{\boldsymbol{y}\in \boldsymbol{E}\ /\ y\ \text{ is a knn of x and x is a knn of y}\}$
- $\forall \boldsymbol{xA}\subset \boldsymbol{E},\ \boldsymbol{ad}_{knn}\left(\boldsymbol{A}\right)=U_{\boldsymbol{x}\in\boldsymbol{A}}\boldsymbol{ad}_{knn}\left(\left\{\boldsymbol{x}\right\}\right)$

Property 2: The knn and k-adjacency pretopologies are of type $\mathcal{V}_{DS}$.

Let $\left(E,ad\right)$ be a pretopological space of type $\mathcal{V}$, we have the following properties:

Proposition 2: $\forall F_1,F_2$ closed subsets of E, $F_1\cap F_2$ is also closed subset.

Proposition 3: $\forall x,y \in E,\ x \neq y$ such that: $F(x) \cap F(y) \neq \varnothing$; then $\forall z \in F(x) \cap F(y)$, $F(z) \subset \left(F(x) \cap F(y)\right)$

Let E be a finite set. Let (E,ad) be a pretopological space of type $\mathcal{V}_D$. We have the following property.

Property 3: The pseudoclosure is the union of the of a subset A is the union of the elementary closed subsets of its elements. $\forall A \in \mathcal{P}(E), ad(A) = U_{x \in A} ad(\{x\})$

The continuity concept, well known in topology, has also been defined in the pretopological framework.

Definition 10 (Continuity) Let (E, ad_E) and (F, ad_F) be two pretopological spaces. Let h be a function defined from (E, ad_E) to (F, ad_F). h is said to be **(m, n) continuous** from E to F if and only if:

$\forall A \in \mathcal{P}(E), h\left(ad_E^m(A)\right) \subset ad_F^n\left(h(A)\right)$. If $m = n$, we said that h is continuous on E.

This notion of continuity defined on a poretopological space will allow to transfer structures between two pretopological spaces. We will use this notion to define a temporal relation allowing to preserve the structures between two consecutive time. We will use this notion to define a temporal relation allowing to preserve structures and properties between two consecutive times.

Let $x \in E$ and A a subset of E. Then $ad(\{x\})$ represents the elements of E "linked" to x, and $ad(A)$ corresponds to the set of elements of E that are "linked" to elements of A, where linked can be interpreted as similar to, influenced by, etc. depending on the application domain. A simple example often used for ad is the neighborhood. Conversely, one can build a pretopological space from a neighborhood relation, or more generally from a binary relation.

PRETOPOLOGY BASED ON A BINARY RELATION

We will present here a way to build a pre-topological space from a binary relation or a family of binary relations (neighborhood family).

Let $\mathfrak{R}$ be a binary relation defined on E. We denote $\mathfrak{R}(x) = \{y \in E / x\ \mathfrak{R}\ y\}$, and we define $\mathfrak{R}^0(x) = \{x\}$, $\mathfrak{R}^1(x) = \mathfrak{R}(x)$ et $\forall p \geq 1$, $\mathfrak{R}^p(x) = \mathfrak{R}\left[\mathfrak{R}^{p-1}(x)\right]$. In a similar way, we denote: $\mathfrak{R}^{-1}(x) = \{y \in E / y\mathfrak{R}x\}$ and $\forall p \geq 1$, $\mathfrak{R}^{-p}(x) = \mathfrak{R}\left[\mathfrak{R}^{-p+1}(x)\right]$. We thus have: $\forall\ A \subset E /\ \mathfrak{R}(A) = U_{x \in A} \mathfrak{R}(x)$.

Based on the binary relation $\mathfrak{R}$, we can define three pretopological spaces (Belmandt, 1993; Brissaud et al., 2011; Hubert Emptoz, 1983) as follows.

Pretopology of p-order ascendants is the pretopological structure on E with the pseudo-closure $ad^{(p)}$ of ascendants of order p defined by:

$\forall A \in \mathcal{P}(E), \quad ad^{(p)}(A) = \{x \in E\ / \exists i, 0 \leq i \leq p, \mathfrak{R}^{-i}(x) \cap A \neq \varnothing\}$

Pretopology of q-order descendants is the pretopological structure on E with the pseudo-closure $ad^{(q)}$ of descendants of order q defined by:

$\forall A \in \mathcal{P}(E), \quad ad^{(q)}(A) = \left\{x \in E \,/\, \exists j, 0 \le j \le q, \mathfrak{R}^j(x) \cap A \neq \varnothing\right\}$

Pretopology of (p,q)-order ascendants-descendants is the pretopological structure on E with the pseudo-closure $ad^{(pq)}$ of ascendants-descendants of order (p,q) defined by:

$\forall A \in \mathcal{P}(E), \quad ad^{(pq)}(A) = \left\{x \in E \,/\, \exists(i,j), 0 \le i \le p, 0 \le j \le p, \mathfrak{R}^{-i}(x) \cap A \neq \varnothing \quad and \ \mathfrak{R}^j(x) \cap A \neq \varnothing\right\}$

Proposition 4: Pretopology of p-order ascendants, q-order descendants and (p,q)-order ascendants-descendants are pretopological spaces of type $\mathcal{V}_D$.

Example: (E, ad_ε) is a pretopological space of 1-order descendants.

TEMPORAL PRETOPOLOGICAL MODEL: A NEW CONCEPT

The notion of continuity enables to perform a pretopological structure transfer from a set E to a set F. This transfer can be used to study subsets of structures changing over time.

In that section, we introduce a new concept named temporal pretopological space that aims to study sub-structures and relationships between elements of those substructures, which both evolve along a temporal dimension $T = \{1, \ldots, n\}$. That notion allows us to establish a generic formalism. Then we present an example of temporal pretopology built from a binary relation. We give two interesting examples of substructures evolving over time. The set E stays the same but the pseudo-closure ad_t may change. Our formalism is based on the definition of a temporal function f_t that must verify the continuity constraint (see definition bellow) between spaces (E, ad_t) and (E, ad_{t+1}).

Definition 11 (Temporal function) f_t is a temporal function between time stamps t and $t+1$ defined on $\mathcal{P}(E)$, if f_t preserves the continuity condition between (E_t, ad_t) and (E_{t+1}, ad_{t+1}), i.e. $\forall A \in \mathcal{P}(E)$, f_t verifies $f_t(ad_t(A)) \subset ad_{t+1}(f_t(A))$.

Definition 12 (Temporal pretopological space) A temporal pretopological space over a sequence of time stamps $T = \{1, \ldots, K\}$ is a sequence of pretopological spaces $\langle (E_1, ad_1), \ldots, (E_K, ad_K) \rangle$ with a temporal function f_t between (E_t, ad_t) and (E_{t+1}, ad_{t+1}) such that $\forall t \in \{1, \ldots, K-1\}, f_t(ad_t(A)) \subset ad_{t+1}(f_t(A))$.

Construction Of Temporal Pretopology Of P-Order Descendants

Let G be a temporal pretopological space (i.e. a sequence of pretopologies) $G = \langle G1, G2, \ldots, GK \rangle$ where $G_i = (E, ad_i)$. We can build a temporal pretopology of p-order descendants from the sequence of pretopologies $(E, ad_i), \forall i \in \{1, \ldots, K\}$.

Definition 3.3 (Temporal relation of descendants) We call temporal relation of descendants the relation $\mathfrak{R}_t$ defined by: $\forall x \in E, \mathfrak{R}_t(\{x\}) = \{y \in E \,/\, y \in ad_{t+1}(\{x\})\}$.

We denote: $\mathfrak{R}_t^0(x) = \{x\}$, $\mathfrak{R}_t^1(x) = \mathfrak{R}_t(\{x\})$ et $\forall p \ge 1$, $\mathfrak{R}_t^p(x) = \mathfrak{R}_t[\mathfrak{R}_t^{p-1}(x)]$.

We generalize the temporal relation of x definition by: $\forall A \in \mathcal{P}(E), \mathfrak{R}_t(A) = U_{x \in A}\ \mathfrak{R}_t(\{x\})$.

Definition 13: (Temporal pretopological space of p-order descendants) We call temporal pretopological space of p-order descendants the pretopological space (E, ad_t^p) whose pseudo-closure is defined as: $\forall A \in \mathcal{P}(E), ad_t^p(A) = \{x \in E \ / \exists i, 0 \leq i \leq p, \mathfrak{R}_t^i(x) \cap A \neq \varnothing\}$.

ad_t^p is called **temporal pseudo-closure of p-order descendants**.

Figure 1. Example of temporal pretopology

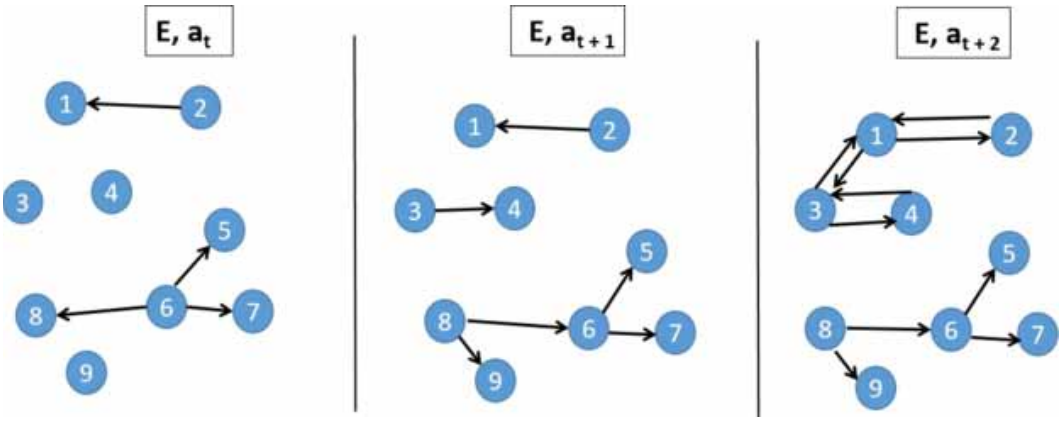


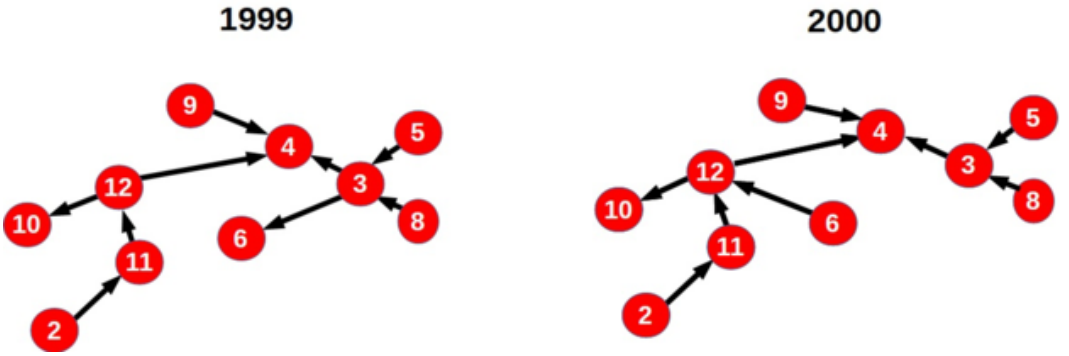
Figure 1 shows an example of temporal pretopological space $G = \{G_t,\ G_{t+1},\ G_{t+2}\}$, where $G_t = (E,\ ad_t)$, $G_{t+1} = (E,\ ad_{t+1})$, $G_{t+2} = (E,\ ad_{t+2})$ and $E = \{1, 2, 3, 4, 5, 6, 7\}$. In that figure, individuals are nodes, an edge $u \rightarrow v$ means that $v \in ad(\{u\})$. Table 1 provides the temporal function of the descendants $\mathfrak{R}_t$ and $\mathfrak{R}_{t+1}$ between t, $t+1$ and $t+1$, $t+2$. Figure 2 shows the Temporal pseudo-closure of descendants of order p=1,2.

The objective of the temporal function is to allow the pseudo-closure to temporally extend up to its closure whilst keeping properties of the pretopological structure G.

Table 1. Temporal function $\mathfrak{R}_t$ between t and $t+1$

x	1	2	3	4	5	6	7	8	9
$\mathfrak{R}_t(x)$	{1}	{1,2}	{3,4}	{4}	{5}	{5,6,7}	{7}	{6,8,9}	{9}
$\mathfrak{R}_{t+1}(x)$	{1,2,3}	{1,2}	{1,3,4}	{3,4}	{5}	{5,6,7}	{7}	{6,8,9}	{9}

Figure 2. Temporal pseudo-closure of descendants of order $\mathbf{p}=1,2$



DEFINITION OF NEW SUBSTRUCTURES EVOLVING OVER TIME

In some application domains, as for example in economics, it is pertinent to analyze sectoral dependencies, such that sectorial influences, which either change over time or remain stable over a period. That allows economists to predict dependencies associated to stable or evolving economic sectors.

In this section, we define two new notions of temporal substructures:

1. Sequence associated to an elementary closed subset and that shows change in influence of a particular element on its p-order descendants.
2. Subset of p-order descendants of an element stable over a fixed period of time k, which we call k-stable substructure.

In the following, $\left(E,ad_t^p\right)$ denotes the temporal pretopological space of p-order descendants where $t\in T=\left\{1,...,n\right\}$.

Evolution of an Elementary Closed Sub-Structure

In the context of structural analysis, closed subsets have often be studied in spaces with a pretopological structure. Moreover, such subsets have been used to build clusters or models in supervised classification. In the following, we define the concept of temporal closed subset associated to an element of the set under study, and propose an algorithm to seek such subsets.

Definition 14 (Temporal evolution of an elementary closed substructure) $\forall x\in E$, we define the temporal evolution of an elementary closed substructure associated to x as the sequence of the elementary closed subsets associated to x at each time stamp t in T. In other words, the sequence $F=<F_1\left(x\right),...F_t\left(x\right),...,F_K\left(x\right)>$ where $F_t\left(x\right)$ is the elementary closed subset associated to x in $\left(E,ad_t\right)$.

F is built iteratively by performing a temporal pseudo-closure at each time stamp t. Figure 3 shows the construction of a sequence of elementary closed subsets associated to $x=8$ by the algorithm 1.

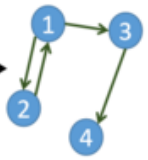
Figure 3. Example of a sequence of temporal closed subsets associated to $x=8$

	1	2	3	4
1	800,71	4189,92	23333,34	194,46
2	1482,37	2000,86	100,69	129,94
3	47,66	55,09	4000,65	42333,91
4	1268,00	2301,97	500,70	14969,40
SUM	3598,732	8547,842	27935,383	57627,701

→

	1	2	3	4
1	0,222	***0,490***	***0,835***	0,003
2	***0,382***	0,168	0,006	0,008
3	0,012	0,005	0,226	***2,693***
4	0,327	0,193	0,028	0,952
MEAN	***0,236***	***0,214***	***0,274***	***0,914***

→



Algorithm 1: Temporal evolution of an elementary closed subset
Input:
- $\left\{\left(E, ad_1\right), \left(E, ad_2\right), ..., \left(E, ad_n\right)\right\}$ sequence of n pretopological spaces
- x an element
Output:
- G^P
G^P = [ ]
For each timestamp t do
$G^P(t) = F_t(x)$
End For
Return G^P

k-Stable Temporal Substructures

A k-stable temporal substructure can be generated iteratively from a 2-stable temporal substructure.

Definition 15 (2-stable temporal substructure) We say that $A \subset E$ is a 2-stable temporal substructure between two consecutive time stamps t and $t+1$, if and only if, $\exists B \subset E$, $A = ad_t\left(B\right) = f_t\left(B\right)$

Definition 16 (k-stable temporal substructure) Let $k \geq 1, A \subset E$ is a temporal substructure k-stable iff $\exists t \in T$, such that A is a 2-stable temporal substructure between t and $t+1$, $t+1$ and $t+2$, ..., $t+k-1$ and $t+k$.

Definition 17 (Maximal k-stable temporal substructure) Let $k \geq 1, A \subset E$ is a maximal k-stable temporal substructure if and only if A is a k-stable temporal substructure and $\nexists B \subset E$ k-stable temporal substructure such that $A \subset B$.

Those substructures can be built iteratively by applying pseudo-closure and intersection operators to an element $x \in E$. Figure 3 shows 2 examples of 3-stable temporal substructures: $\{1,2\}$ and $\{5,6,7\}$; and $\{5,6,7,8,9\}$ is a maximal 2-stable temporal substructure.

Algorithm 2: k-stable temporal substructure
Input:
- $\left\{\left(E, ad_1\right), \left(E, ad_2\right), ... , \left(E, ad_n\right)\right\}$ a sequence of n pretopological spaces
- x an element
Output:
- G^P
G^P = [ ]
i=0
For each pretopological space do
If first space then

```
                    Gᵖ(i)= ℜᵢᵖ (x)Ç adᵢᵖ (x)
                    //search of 2-stabes subset between (E, adᵢ) and
(E, adᵢ₊₁)
          Else
                     Gᵖ(i)= Find 2-stable subset between Gᵖ(i-1) and
(E, adᵢ)
          End If
          i = i+1
End For
Return Gᵖ
```

EXPERIMENTAL RESULTS

Experiments were performed on several data sets: 2 data sets from the Stanford Network Analysis Platform (SNAP), and 2 real-word data sets related to sector influence in New-Caledonia economics (data set 1) and in Metropolitan France economics (data set 2) respectively. Table 2 gives a brief description of those data sets.

COLLEGEMSG[1] is a data set containing private messages sent through an online social network of the University of California. Edge $u \rightarrow v$ means that user u sent a private message to user v.

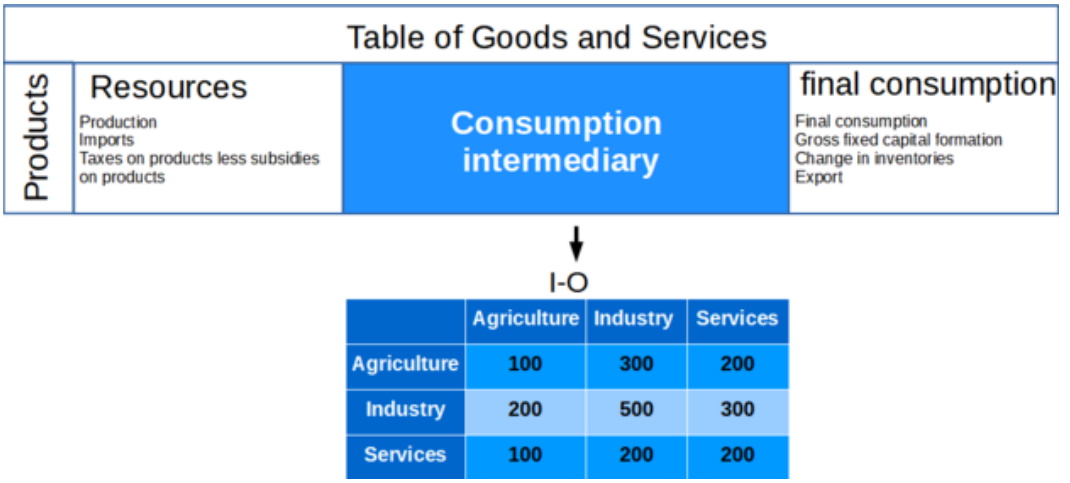
Table 2. Description of data

	NB instances	Temporality	Time stamps
Economic data 1	204	Year	17
Economic data 2	1139	Year	67
College message	1899	Month	9
Email-UE	986	Day	12

EMAIL-EU[2] is a set of emails exchanged by members of an European Research Institution. Edge $u \rightarrow v$ means that individual u sent an email to individual v.

Figure 4 shows examples of the temporal evolution of some individuals for the 2 data sets COLLEGEMSG and Email-UE. The algorithm highlights the evolution of interactions among a group of persons. We can observe that individual 3 from COLLEGEMSG data set interacts indirectly with individual 155 at some stage and directly some time later. Same remark for individual 194 interactions with individual 311 in the EMAIL-EU data set.

Figure 4. Examples of elementary temporal closed subsets



Application to Sectoral Econometrics

The economic activity of a country can be modeled by various monetary flows observed on an annual basis. To represent the state of their national accountings, European countries use tables displaying those monetary flows according to the European System of Accounts (ESA). In that system, the input-output table (IOT) is used to ensure the consistency of national accounts. It describes and summarizes operations on goods and services with regard to products and industries per annum. In addition, IOT is used to analyze inputs-outputs in order to model sector interdependence.

Inputs and outputs are consumed and sold products by sectors, in monetary value. An industry frequently uses inputs produced by other industries. Similarly, the production of that industry can serve as inputs for other industries. A particularly simple example, provided by W. Leontief in (Leontief, 1986), will enable us to illustrate how to construct an input-output table from goods and services accounts represented by 3 blocks (Figure 5).

Figure 5. Input-output table

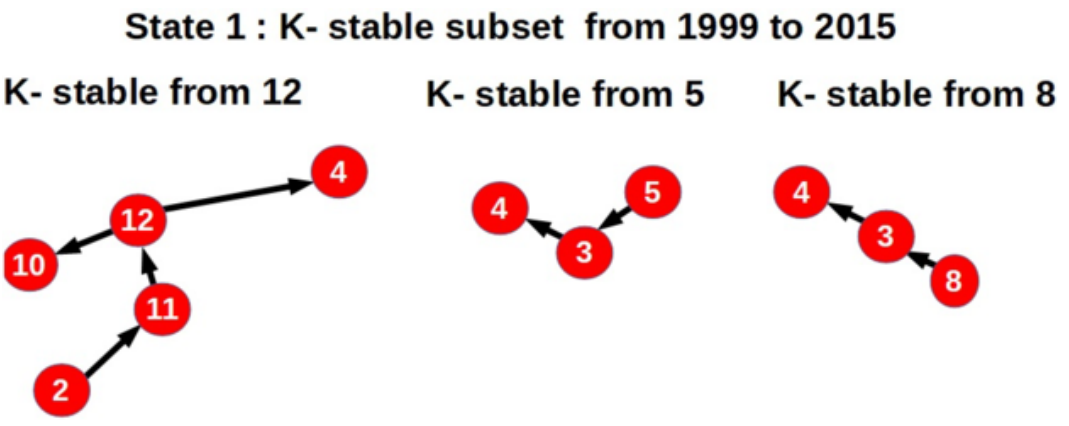
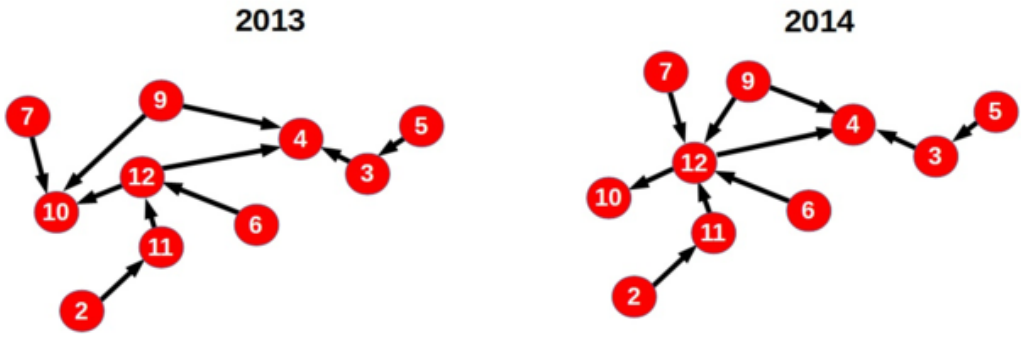


Figure 6. Example of setting of the I/O table



The following study deals with the input-output tables of New Caledonia per annum between 1999 and 2015. Each IOT contains the 12 industries defined in Table 3.

According to Figure 6, the 17 tables have been normalized. We recall here the way pretopologies of descendants were generated from the data. For each sector, the annual average threshold enables to decide whether a sector k impacts another sector j. This influence is taken into consideration if the value in the normalized table $A\left(j,k\right)>s_k$, where s_k is the average threshold of the industry k at year n. An oriented graph is built for each year n, based on the thresholds for that year. If the normalized intermediate consumption of a sector k comes from another sector j with a value greater than the threshold (of sector k), then an edge j k is added to the graph.

Table 3. Industries in New Caledonia

Industry	Vertex number
Agriculture, hunting, forestry, fishing, lifestock farming	1
Food industry	2
Nickel industry	3
Other industries	4
Energy	5
Building and public works	6
Trade	7
Transports and telecommunications	8
Banks and insurances	9
Services mainly towards companies	10
Services mainly towards households	11
Administration	12

We applied the algorithm on each vertex in the pretopology of descendants representing the normalized IOT from 1999 to 2015. For example, in Figure 7 and Figure 8[REMOVED REF FIELD], by applying successively the pseudo-closure (from 1999 up to 2014), one can find the pattern $\left\{11,12,10\right\}$ in all IOTs.

According to domain experts, the proposed method enables to highlight an evolution of New Caledonia economics characterized by the development of a local processing industry over the two last decades.

Expert's Interpretation

At the end of the 1990's and at the beginning of the 2000's (cf. Figure 7), "other industries" (4) were mainly oriented towards mining (3) and administration (12). That corresponded to either subcontracting mining activities or technical activities such as maintenance or else for administrations. The bank sector (9) also benefitted from (4).

Next, the sector of "other industries" (4) has grown significantly thanks to, on one hand, the construction of 2 world-class metallurgical sites (Vale NC and Koniambo Nickel), and on the other hand a political willingness to create a local industry to substitute imports. That political orientation was accompanied by tariff barriers, even by quotas, in order to favor the local processing industry which products are more expensive than imported ones. The political aim also intended to create numerous jobs in New Caledonia.

That explain why, ten years later (cf. Figure 8), production of other sectors such as trade (7) could then depend on diverse local industries (4) (year 2014). On the contrary, the later could depend on the production of sectors such as the Building and Public Works sector (6) (year 2013).

Figure 7. (a) Examples of temporal closed elementary subsets

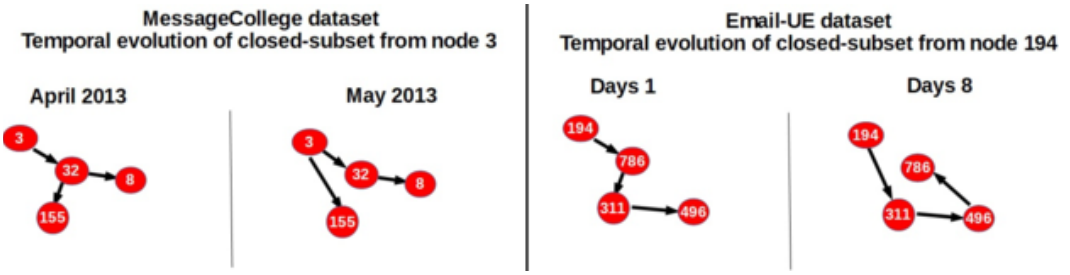


Figure 8. (b) Examples of temporal closed elementary subsets

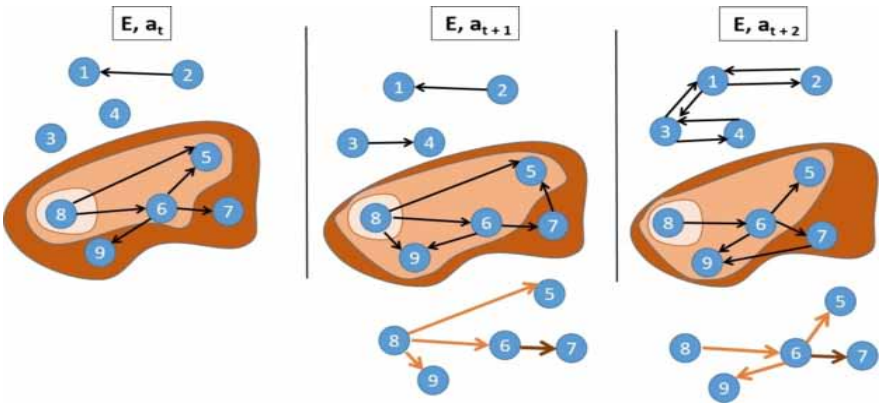
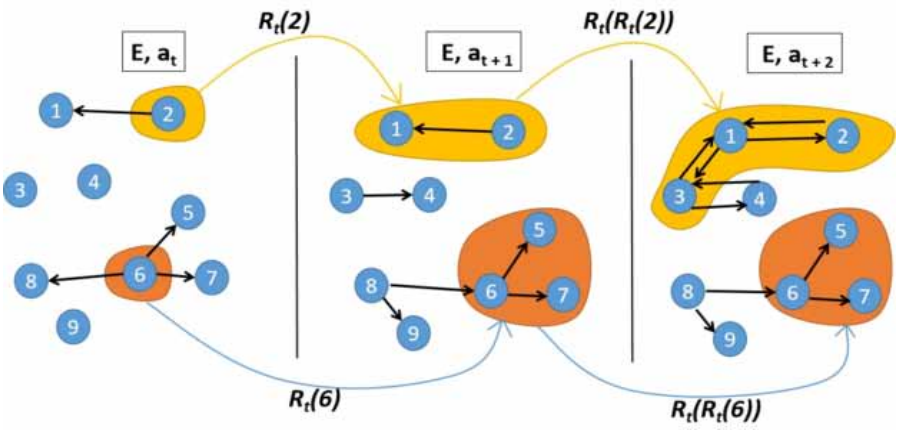


Figure 9[REMOVED REF FIELD] shows examples of maximal k-stable substructures from 3 different sectors. Sector 2 (food industry) influenced sectors 11 (services mainly towards households), 12 (administration), 10 (services mainly towards companies), and 4 (other industries) between 1999 and 2015. In addition, we can notice that sector 4 is indirectly impacted by the 3 sectors involved in the example of Figure 9.

Figure 9. k-stables substructures



CONCLUSION AND PERSPECTIVES

We have established a first formulation of the concept of temporal pretopological space. More precisely, from an example of temporal relation, we have defined the temporal pretopological space of p-order descendants. We introduced two new notions of temporal substructures, generated by applying a temporal pseudo-closure, and that could be extracted to analyze data from a structural point of view.

The present works opens numerous perspectives. One could be to optimize the two algorithms before applying them to large data sets. Another direction could be to study the properties of temporal pretopological spaces. A further objective could be to combine structural analysis of individual influences and characteristics describing individuals, which might also change over time. Exploration of substructures combined to pattern extraction (especially itemsets) could then lead to a cross analysis. In the example of economic data, an important point for economists is the impact of the influence threshold between sectors. It could thus be pertinent to define temporal pretopological structures with weighted influences instead of keeping only influences greater than a threshold.

Finally, in the present work, temporal pretopology is generated from an example of binary relation. Another interesting approach would be to define other criteria enabling to search for substructures in data issued from a family of binary relations or family of neighborhoods.

FUNDING AGENCY

Publisher has waived the Open Access publishing fee.

ACKNOWLEDGMENT

This research was supported by the CNRT and the PIL.

Piegay, E., Selmaoui, N., & Leschi, C. (1995). Crest lines detection by valleys spreading. *International Conference on Computer Analysis of Images and Patterns*, 270–277. doi:10.1007/3-540-60268-2_306

Selmaoui, N., Leschi, C., & Emptoz, H. (1993). Crest Lines Detection in Grey Level Images: Studies of Different Approaches and Proposition of a New One. *CAIP'93*. *LNCS*, *719*, 157–164. doi:10.1007/3-540-57233-3_21

REFERENCES

Arnaud, G. M., Lamure, M., Terrenoire, M., & Tounissoux, D. (1986). Analysis of the connectivity of an object in a binary image: a pretopological approach. *Proc. of the 8th IAPR Conference*.

Auray, J.-P., Duru, G., & Mougeot, M. (1979). A pre-topological analysis of the input-output model. *Economics Letters*, *2*(4), 343–347. doi:10.1016/0165-1765(79)90048-X

Basileu, C., Ben Amor, S., Bui, M., & Lamure, M. (2012). Prétopologie stochastique et réseaux complexes. *Stud. Inform. Univ.*, *10*(2), 73–138.

Belmandt, Z. (1993). Manuel de prétopologie et ses applications: Sciences humaines et sociales, réseaux, jeux, reconnaissance des formes, processus et modèles, classification, imagerie, mathématiques. *Hermes*.

Brissaud, M. (1971). Topologie et prétopologie. *Publications Econométriques*, *4*, 254–268.

Brissaud, M. (1975). Les esapces prétoplogiques. *Compte-Rendu de l'académie Des Sciences, 280*(A), 705–708.

Brissaud, M., Lamure, M., Milan, J.-J., Auray, J.-P., Nicoloyannis, N., Duru, G., Terrenoire, M., Tounissoux, D., Zighed, D. A., Bonnevay, S., Van Le, T., Bui, M., Ben Amor, S., Levorato, V., & Kabachi, N. (2011). *Basics of pretopology*. Hermann. https://hal.archives-ouvertes.fr/hal-00605389

Bui, Q. V. (2018). *Pretopology and Topic Modeling for Complex Systems Analysis : Application on Document Classification and Complex Network Analysis*. http://www.theses.fr/2018PSLEP034

Cleuziou, G., Buscaldi, D., Levorato, V., & Dias, G. (2011). A pretopological framework for the automatic construction of lexical-semantic structures from texts. *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, 2453–2456. doi:10.1145/2063576.2063990

Dalud-Vincent, M., Brissaud, M., & Lamure, M. (2001). Pretopology as an extension of graph theory : The case of strong connectivity. *Journal of Applied Mathematics*, *5*(4), 455–472. https://halshs.archives-ouvertes.fr/halshs-00470194

Duru, G. (1980). *Contribution à l'étude des structures des systèmes complexes dans les sciences humaines*. Academic Press.

Emptoz, H. (1983). *Modèle prétopologique pour la reconnaissance des formes application en neurophysiologie*. INSA DE LYON.

Emptoz, H., Hashom, A., & Terrenoire, M. (1981). *Fonctions structurantes et classification automatique*. Congrès AFCET.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). Density-based spatial clustering of applications with noise. *Int. Conf. Data Mining and Knowledge Discovery*, *240*, 6.

Lamure, M. (1987). *Espaces abstraits et reconnaissance de formes. Application au traitement des images digitales*. Univ. Claude Bernard Lyon 1.

Le, T. V., Kabachi, N., & Lamure, M. (2007). A clustering method associated pretopological concepts and k-means algorithm. In *Recent advances in stochastic modeling and data analysis* (pp. 529–536). World Scientific. doi:10.1142/9789812709691_0062

Leontief, W. (1986). *Input-output economics*. Oxford University Press.

Levorato, V. (2011). *Modeling groups in social networks*. Academic Press.

Piegay, E. (1997). *Groupement, multirésolution, prétopologie : analogies entre la segmentation d'images et la classification automatique*. INSA de Lyon.ENDNOTES

[1] https://snap.stanford.edu/data/CollegeMsg.html
[2] https://snap.stanford.edu/data/email-Eu-core.html

Nazha Selmaoui-Folcher is Associate Professor (HDR) at the University of New Caledonia since 1998. She received her Ph.D. degree in 1992 from the "Institute National des Sciences Appliques" at Lyon (France) and her Habilitation degree in 2012 from the University of Lyon I. Her current research interests are image analysis, machine learning, artificial intelligence, graph mining, data mining on temporal series of satellite images, spatio-temporal data and its application to health and environmental sciences. She is involved in the program committees of many data mining and machine learning conferences and Journals. She was the leader of a multidisciplinary laboratory on matter and environment 2012-2016 (PPME). She was the coordinator for the ANR FOSTER projects dedicated to Knowledge Discovery on spatio-temporal Databases and Application to Soil Erosion (2011-2014). She participates to ANR project SpiRAL (Soils, rainfall and leptospirosis Understanding leptospirosis environmental contamination). She is member of LABEX-CORAIL (laboratory of excellence, Funded by French Agency) and associate member of IPAL (International Research Lab in Singapore).

Samuel Gorohouna is an Associate Professor in Economics at the University of New Caledonia and a researcher at the Legal and Economic Research Laboratory (LARJE). Inequalities in the labour market and access to housing constitute his key area of research. He is also involved in work on integrating indigenous communities into the market economy, on greening the tax system and on voting behaviour. He teaches microeconomics and development economics, with a focus on the growth of small regional economies. He is Vice-President of the University and Head of the Baco Campus in Koné.

Claire Leschi is currently Associate Professor at the National Institute of Applied Sciences of Lyon. She is a member of the LIRIS Laboratory. Her research interests are machine learning, artificial intelligence and data mining on spatio-temporal data, with applications in education and environmental sciences.

Catherine Ris is Professor of Economics at the University of New Caledonia, Director of LARJE (Research Centre on Economics and Law). She received her PhD from the University of Lyon and was a post-doctoral fellow at the Research Centre for Education and the Labour Market (ROA) at Maastricht University. She returned to UNC in 2014 after one year as a Research Fellow at the Centre for Development Studies, University of Auckland. Her main research interests and major academic works are in development economics, public policy evaluation, economic and social inequality, Pacific development, quantitative and qualitative data analysis.