Technological Forecasting & Social Change 111 (2016) 110–123

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Technology forecasting in the National Research and Education Network technology domain using context sensitive Data Fusion

L. Staphorst [a,b,1,2], L. Pretorius [b,2], M.W. Pretorius [b,2]

[a] *South African National Research Network (SANReN), Council for Scientific and Industrial Research, South Africa*
[b] *Department of Engineering and Technology Management, University of Pretoria, South Africa*

ARTICLE INFO

Article history:
Received 31 December 2013
Received in revised form 2 June 2016
Accepted 2 June 2016
Available online 1 July 2016

Keywords:
Technology intelligence
Technology indicators
Technology forecasting
Data Fusion
Structural Equation Modeling
National Research and Education Network

ABSTRACT

Using inductive reasoning this paper develops a framework for the Structural Equation Modeling based context sensitive Data Fusion of technology indicators in order to produce Technology Forecasting output metrics. Data Fusion is a formal framework that defines tools, as well as the application of these tools, for the unification of data originating from diverse sources. Context sensitive Data Fusion techniques refine the generated knowledge using the characteristics of exogenous context related variables, which in the proposed framework entails non-technology related metrics. Structural Equation Modeling, which is a statistical technique capable of evaluating complex hierarchical dependencies between latent and observed constructs, has been shown to be effective in implementing context sensitive Data Fusion. For illustrative purposes an example model instantiation of the proposed framework is constructed for the case of the National Research and Education Network technology domain using knowledge gained through action research in the South African National Research Network, hypotheses from peer-reviewed literature and insights from the Trans-European Research and Education Network Association's annual compendiums for National Research and Education Network infrastructure and services trends. This example model instantiation hypothesizes that a National Research and Education Network's infrastructure and advanced services capabilities are positively related to one another, as well as to the contextual influence it experiences through government control. Also, positive relationships are hypothesized between a National Research and Education Network's infrastructure and advanced services capabilities and its usage, which is defined as the technology forecasting output metric of interest for this example. Data from the 2011 Trans-European Research and Education Network Association compendium is used in the Partial Least Square regression analysis of the example model instantiation, which confirms all hypothesized relationships, except the postulation that a National Research and Education Network's infrastructure and advanced services capabilities are positively related. This latter finding is explained by observing the prevalence of technology leapfrogging in the National Research and Education Network global community.

1. Introduction

Technological advancement continues at an astounding rate, seemingly following exponential growth models such as Moore's (Mack, 2011), Nielsen's Law (Nielsen, 1988) and Metcalfe's Law (Metcalfe, 1995). Driven not only by the invention, innovation and diffusion of new technologies, but also by the move to the paradigms of globalization and open innovation (Nyberg and Palmgren, 2011), this has created highly competitive global markets for technology based products and services (Porter, 2007). Hence, the survival, growth and profitability of firms that play in these markets depend highly on their ability to monitor current, as well as predict future technological changes in order to create a solid and sustainable technological base that can withstand, or adapt to rapidly changing market requirements (Porter, 2007). Moreover, firms need to effectively and efficiently manage technological changes both internally and externally if they are to create sustainable competitive advantages in rapidly high-tech markets (Lichtenthaler, 2004). Technology Intelligence (TI), which is a core process within the discipline of technology management, involves the process of capturing technology related data, converting this data into information by determining relational connections and refining information to produce knowledge that can guide strategic decision makers during strategic planning (Lichtenthaler, 2004; Chang et al., 2008). Technology indicators, such as technology maturity and degree of innovation, are those measureable sources of technology related data that allow for the direct characterization and evaluation of technologies over their whole life

---

E-mail addresses: leon.staphorst@gmail.com (L. Staphorst), leon.pretorius@up.ac.za (L. Pretorius), tinus.pretorius@up.ac.za (M.W. Pretorius).

[1] Postal Address: PO Box 395, Pretoria, 0001, South Africa.

[2] Postal Address: Department of Engineering and Technology Management, University of Pretoria, Lynnwood Road, Hatfield, Pretoria, 0001, South Africa.

http://dx.doi.org/10.1016/j.techfore.2016.06.012
0040-1625/© 2016 Elsevier Inc. All rights reserved.

cycle (Chang et al., 2008). Scrutinizing the information that has been distilled from a set of technology indicators in a forward-looking approach, commonly referred to as Future-oriented Technology Analysis (FTA), can potentially provide decision makers with Technology Forecasting (TF) knowledge, amongst others (Porter, 2005).

Buchroithner (Buchroithner, 1998) and Wald (Wald, 1997) define Data Fusion (DF), which was developed in the military domain for the generation of quality tactical knowledge through the multi-layered processing of sensor data (Wald, 1999), as "… a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of 'greater quality' will depend upon the application." Within the discipline of DF, context can be viewed as synonymous with a situation, which in turn is defined as a set of relational connections (i.e. instantiated relations) (Steinberg, 2009). Context can be used in each level of the DF process in order to refine data alignment and association, as well as during situation state estimation (Steinberg, 2009). Recently, context sensitive DF techniques have been explored which effectively refine the generated knowledge at each level of processing based on the characteristics of exogenous context-related variables (Steinberg, 2009).

Regression analysis constitutes a family of statistical techniques geared at modeling and analyzing the relationship between dependent and independent variables from empirical data (Haenlein and Kaplan, 2004). Moreover, regression analysis attempts to explain the variations in dependent variables as functions (commonly referred to regression functions) of variations in independent variables (Haenlein and Kaplan, 2004). With this knowledge it is then possible to perform prediction and forecasting of the values that dependent variable will assume for specific independent variable values (Haenlein and Kaplan, 2004). Classic regression techniques (such as multiple regression, discriminant analysis, logistic regression and analysis of variance) can be classified as first generation techniques, since these techniques explicitly assume independence between multiple dependent variables (Haenlein and Kaplan, 2004). This, unfortunately, limits the ability of such techniques to comprehensively model complex interrelationships, such as the interplay between two or more output variables in a TF model. More specifically, classic first generation regression techniques are not able to model the potential mediating or moderating effect that output variables could have on one another. To overcome this limitation, Jöreskog (Jöreskog, 1973) proposed covariance based Structural Equation Modeling (SEM) as a second-generation technique, which allows for the simultaneous modeling of relationships amongst multiple dependent and independent constructs. A further inherent limitation of first generation regression techniques is their explicit assumption that all dependent and independent variables are directly observable (Haenlein and Kaplan, 2004). This assumption implies that all variables' values can be directly obtained from real-world sampling experiments (Haenlein and Kaplan, 2004). As such, any variables that cannot be directly observed need to be considered unobservable and have to be excluded from first generation regression models (Haenlein and Kaplan, 2004). However, such unobservable variables, commonly referred to as latent constructs, are supported by SEM. Steinberg postulated that SEM is ideally suited to implement context sensitive DF (Steinberg, 2009; Steinberg and Rogova, 2008). Not only does SEM support the complex structural models used in situation state estimation (as is required in TF), it also allows for non-linear and non-Gaussian factors and cyclical dependencies amongst model variables that can be either latent or directly observable (Steinberg, 2009).

According to Sohn and Moon (Sohn and Moon, 2003) most TF techniques rarely take into account the structural relationships amongst technology indicators and TF output metrics. SEM, however, provides an advantage over these limited TF techniques by allowing for the modeling of complex hierarchical relationships between technology indicators and TF outputs metrics. Sohn and Moon (Sohn and Moon, 2003) have shown that SEM, which can be viewed as a generalization of factor and path analysis methods such as Bayesian Networks (Steinberg, 2009), can successfully implementing TF of the Technology Commercialization Success Index (TCSI) TF output metric.

An NREN is a specialized broadband network connectivity and service provider that explicitly caters for the needs of the research and education communities of a country (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012). In some instances, NRENs also service the needs of other public sector entities, such as hospitals, municipalities and libraries. Typically, one NREN is present per country (for example SANReN (The South African National Research Network (SANReN), 2013) in South Africa and the Joint Academic Network (JANET) in the United Kingdom), although separate NREN entities could potentially exist to service distinct in-country research and education sectors or geographic areas, for example the Energy Sciences Network (ESnet) and Kansas Research and Education Network (KanREN) in the United States (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012). NREN'si are built primarily on fiber optic cabling infrastructure and provide researchers, educators and students with unparalleled connectivity speeds and advanced services at a fraction of the price of commercial network providers (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012). These networks are currently experiencing rapid technology driven changes, resulting in evolving business models, innovative infrastructure solutions and service offerings, as well as increased international collaboration (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011).

The objectives of this paper are twofold: Firstly, the paper builds on the work of Steinberg (Steinberg, 2009; Steinberg and Rogova, 2008), as well as Sohn and Moon (Sohn and Moon, 2003), by proposing a framework for the SEM based DF of technology indicators in order to produce TF output metrics. The proposed framework is an evolved and improved version of the framework first proposed in (Staphorst et al., 2013). Secondly, application of the proposed framework is illustrated by through the use of a model instantiation example in the NREN technology domain. The proposed NREN example model instantiation was constructed using insights gained through action research in the South African National Research Network (SANReN) (The South African National Research Network (SANReN), 2013), insights from Trans-European Research and Education Network Association's (TERENA) NREN compendium for 2012 on global NREN infrastructure and services trends (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012), as well as hypotheses and postulations in peer-reviewed literature. It is important to view the example NREN model instantiation presented in this paper as a mere illustrative example of the use of the proposed framework, not as a definitive platform for TF in the NREN domain.

The paper is structured as follows: Firstly, a theory review is presented on the use of SEM for context sensitive DF and the use of SEM in TF, as well as taxonomy of technology indicators and forecasting output metrics. An evolved version of the framework for SEM based DF for TF proposed by Staphorst, Pretorius and Pretorius in (Staphorst et al., 2013) is then developed through inductive reasoning. The example NREN model instantiation of the framework is then presented, including the definition of a number of research propositions relevant to this example model instantiation. This is followed by a quantitative evaluation of the example model instantiation using cross-sectional data extracted from TERENA's NREN compendiums for 2011 (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011), including an evaluation of the research propositions defined for the example model instantiation. Lastly, the paper presents an evaluation of the reliability and the validity of the example NREN model instantiation.

2. Theory and framework development

Steinberg postulated that SEM is one potential statistical tool that lends itself naturally to implement DF, with the added benefit that it allows for the inclusion of context sensitivity during the solving of DF inferencing problems (Steinberg, 2009; Steinberg and Rogova, 2008; Sohn and Moon, 2003). Sohn and Moon showed in (Sohn and Moon, 2003) that SEM can be used as a regression technique to evaluate a multi-layered hierarchal model through progressive aggregations and refinements of input technology indicator data in order to produce a reliable statistical estimate of a TF output metric (Staphorst et al., 2013). In the following subsections it will be shown that, through inductive reasoning, Soon and Moon's (Sohn and Moon, 2003) use of SEM for TF and Steinberg's use of SEM to implement context sensitive DF (Steinberg, 2009) can be combined to develop a framework for SEM based DF for TF. This was first proposed by Staphorst, Pretorius and Pretorius in (Staphorst et al., 2013).

2.1. Structural Equation Modeling for context sensitive Data Fusion

Within SEM theory distinction is made between exogenous and endogenous latent constructs, with the former being variables that are not explained by the internal interrelationships embodied by the model and therefore always act as independent variables (Haenlein and Kaplan, 2004). Due to its generality, SEM terminology does not refer to regression analysis' dependent and independent variables, but rather only to exogenous constructs, which are independent variables that are not functions of any relationship in the model, as well as endogenous constructs, which are either dependent or independent variables that are explained by the relationships with other dependent and/or independent variables present in the model.

With reference to the indicators measured as proxies to represent latent constructs, such latent constructs can be classified as follows (Haenlein and Kaplan, 2004): A latent construct with reflective indicators is one in which all measured indicator proxies, also commonly referred to as factors, are expected to have high correlations to the latent construct, as well as other potential reflective indicators, Therefore it will have the ability to represent the variance in the unobserved variable sufficiently. In contrast, latent constructs with formative indicators are those that are represented by a weighted combination of indicators that are not highly correlated to either the latent construct itself, or the other formative indicators included in the weighted combination. The formative indicators of a latent construct can therefore be seen as representing different dimensions of this construct.

Although Jöreskog in 1973 (Jöreskog, 1973) originally proposed that the parameters of a SEM model be estimated using covariance based techniques, of which the LISREL program that was developed by Jöreskog in 1975 is arguably the most popular, variance based techniques, also commonly referred to as component based techniques, have also gained popularity (Haenlein and Kaplan, 2004). Partial Least Squares (PLS), which was first introduced by Wold (Wold, 1975) as Non-linear Iterative Partial Least Squares (NIPALS), is one such variance based technique (Haenlein and Kaplan, 2004). While covariance based techniques attempt to minimize the difference between the sample covariance values and those predicted by the regression model, which is equivalent to estimating the model parameters such that the covariance matrix of the observed measurements is reproduced, PLS regression, which is also sometimes referred to Projections to Latent Structures, focuses on maximizing the variance of the dependent variables explained by the independent variables (Haenlein and Kaplan, 2004).

It is common practice in SEM to represent models using path diagrams that depict the exogenous and endogenous constructs, the path coefficients of interconnections between these constructs, reflective and formative indicators, as well as the loadings of these indicators on constructs (Staphorst et al., 2013). Graphically any SEM path diagram can be condensed into layered groupings of constructs, as is shown in Fig. 1 (Staphorst et al., 2013). While the groupings and layering of these groupings can be done based on any arbitrary set of criteria, this paper's proposed of use of SEM to implement context sensitive DF for TF warrants distinct groupings of exogenous and endogenous constructs, representing context and technology related constructs, respectively. Furthermore, it is proposed that the layering of the groupings is performed in such a manner that context and technology related constructs of a similar nature or complexity occupy the same DF layer.

Recall that DF is essentially a framework for the multi-layered refinement of estimates of problem variables from multiple measurements, either directly or indirectly observable (Steinberg, 2009). The Joint Directors of Laboratories' Data Fusion Group (JDL/DFG) recognized that, in a military environment, DF entails the progressive aggregation and refinement of sensor data in order to produce quality tactical knowledge (Steinberg, 2009). In an attempt to standardize the structure of the multi-layered DF process across all possible military applications and implementations, the JDL/DFG defined the following six levels of processing (Steinberg, 2009): Level 0 involves signal/feature/subject assessment, level 1 involves object assessment, level 2 involves situation assessment, level 3 involves impact assessment, level 4 involves process refinement and level 5 involves user refinement.

While these JDL/DFG DF level definitions are not appropriate for the use of DF in TF, the concept of multi-layered progressive aggregation and refinement of measurement indicator data is core to this paper's proposed framework. Hence, Fig. 2 depicts a more generalized context sensitive DF framework that supports *N* levels of aggregation, with *N* an integer that represents the user of the framework's required number of levels of aggregation and refinement. Applicable context related information can be incorporated into the aggregation and refinement process at any of the *N* levels of DF processing.

By noting that SEM is capable of the simultaneous modeling of relationships amongst multiple dependent and independent constructs, Steinberg (Steinberg, 2009) postulated that SEM is one potential statistical tool that lends itself naturally to implement DF. Moreover, based on the following argumentation Steinberg (Steinberg, 2009) showed that SEM allows for the inclusion of context sensitivity during the solving of DF inferencing problems: Firstly, Steinberg (Steinberg, 2009) defined a situation, or a context, as a set of relationships, where a relationship can be viewed as a specific instantiated relation. In general, context is used in DF inferencing problems in order to refine ambiguous estimates, explain available data and constraint processing during data acquisition, cueing or fusion (Steinberg, 2009). Next, Steinberg harmonized DF and SEM terminology by noting that DF problem variables are in fact SEM endogenous constructs, context variables can be viewed as SEM exogenous constructs and classic DF sensor measurements are the reflective and formative indicators present in SEM.

2.2. Technology indicators and forecasting output metrics

According to Porter and Cunningham (Porter and Cunningham, 2004) technology indicators employ empirical information to estimate technology characteristics that affect technological advance and successful commercialization. Watts and Porter (Watts and Porter, 1997) state that technology indicators are empirical measures stemming from models of technological innovation and progression, such as the S-curve. Nyberg and Palmgren (Nyberg and Palmgren, 2011) expands on these definitions by describing technological indicators as those indices or statistical data that allow for the direct characterization of characteristics of technology throughout their life cycles in order to allow decision makers to take strategic actions. Such indicators can in general be divided into three major categories based on their intended function: input indicators, byput indicators and output indicators (Nyberg and Palmgren, 2011; Grupp, 1998). Grupp (Grupp, 1998) states that input indicators are variables related to drivers of technological progress, byput indicators are variables that are related to sub-phenomena of the technological progress and output indicators are variables related to the qualitative, quantitative or value-rated progress in process or product development (Nyberg and Palmgren, 2011).

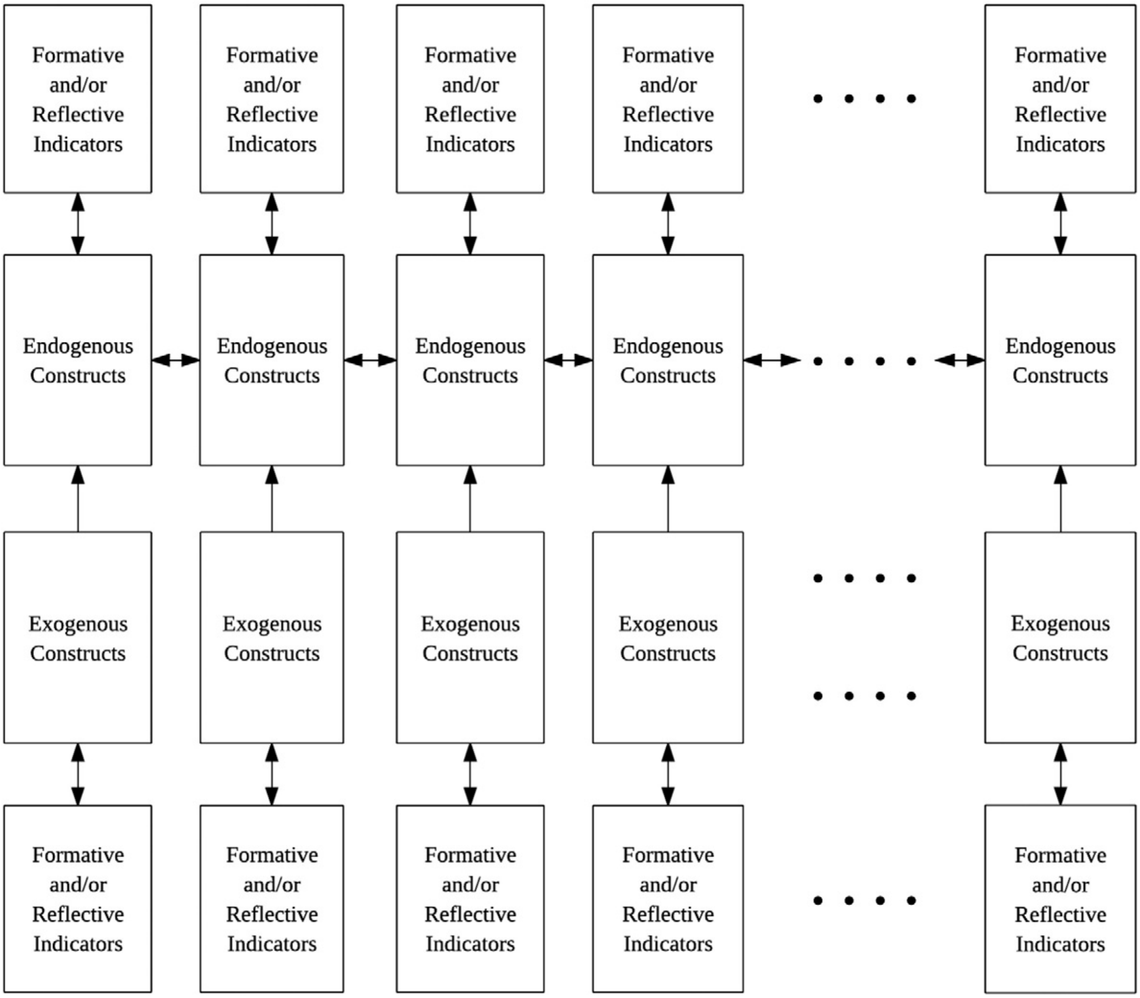

Fig. 1. Grouping and layering of constructs in a SEM path diagram.

As depicted in Fig. 3 input, byput and output technology indicators can be grouped based on similar characteristics or complexity. Furthermore, complex relationships can exist between input, byput and output technology indicators. For the purposes of this paper it is assumed that input technology indicators drive technology processes at Level 0, byput technology indicators indicate sub-phenomena between Level x-1 and Level x for $x = 1, 2, 3, \ldots, N$-1, and output technology indicators indicate product or process related progress at Level x for $x = 1, 2, 3, \ldots, N$-1.

A wide variety of sources exist that can be used to harvest technology indicators, ranging from patent databases and scientific publications (Porter and Cunningham, 2004), through to the rumor mill and financial market indicators (Nyberg and Palmgren, 2011). In monitoring and mining these potential sources of technical indicators, bibliometrics have emerged as one of the most popular set of quantitative techniques (Nyberg and Palmgren, 2011). Bibliometrics uses counts of citations, publications or patents to produce indicators of technological progress in a specific technology domain (Nyberg and Palmgren, 2011).

Various frameworks have been proposed for the systematic categorization of technology indicators. In (Nyberg and Palmgren, 2011) Nyberg and Palmgren presents a succinct summary of the frameworks proposed by Watts and Porter (Watts and Porter, 1997), Grupp (Grupp, 1998) and Chang (Chang, 2007), which is repeated here: The Watts and Porter (Watts and Porter, 1997) framework consists of the following three categories:

- *Technology Life Cycle Status Indicators*: Based primarily on the S-curve, these metrics determine the level of progress of a technological development along its life cycle, as well as the growth rate of the technology (Nyberg and Palmgren, 2011).
- *Innovation Context Receptivity Indicators*: These indicators gauge the sufficiency of supporting technology, as well as the development of standards and regulations surrounding the technology under investigation (Nyberg and Palmgren, 2011).
- *Market Prospect Indicators*: The potential commercial payoffs of the technology are considered by this type of indicator. Of specific importance with these indicators are factors such as technology application areas, intellectual property and market competitiveness (Nyberg and Palmgren, 2011).

Although Grupp was the instigator of the general function based classification of technology indicators into input, byput and output indicators (Grupp, 1998), he originally referred to these three types of indicators based on the stage in the technology's life cycle at which the measurement was performed:

- *Resource Indicators*: This input indicator type measures the various possible expenditures on research, development and innovation (Nyberg and Palmgren, 2011).
- *Research and Development* (*R&D*) *Results Indicators*: This is the output indicator type, which measures qualitative, quantitative, or value rated advances in production processes or products (Nyberg and Palmgren, 2011).
- *Progress Indicators*: Indicators of this type, for example the

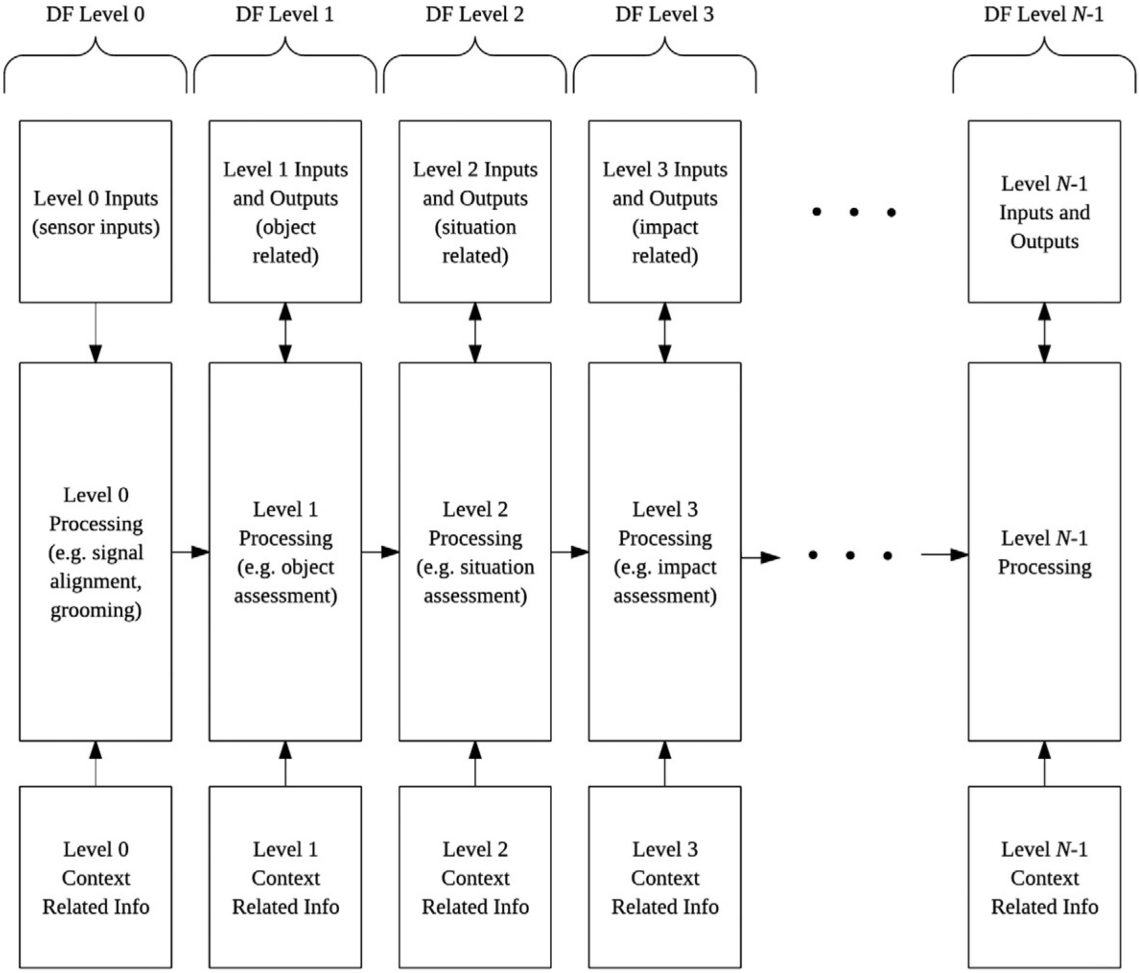

Fig. 2. Generalized framework for context sensitive DF.

technometric indicator (Grupp, 1998) that measures the number of features or product specifications, are byput metrics that measures sub-phenomena of the technological progress (Nyberg and Palmgren, 2011).

The Technology Indicator Ontology (TIO) proposed by Chang (Chang, 2007) divides technology indicators into the following two broad groupings, each with a number of sub-groups:

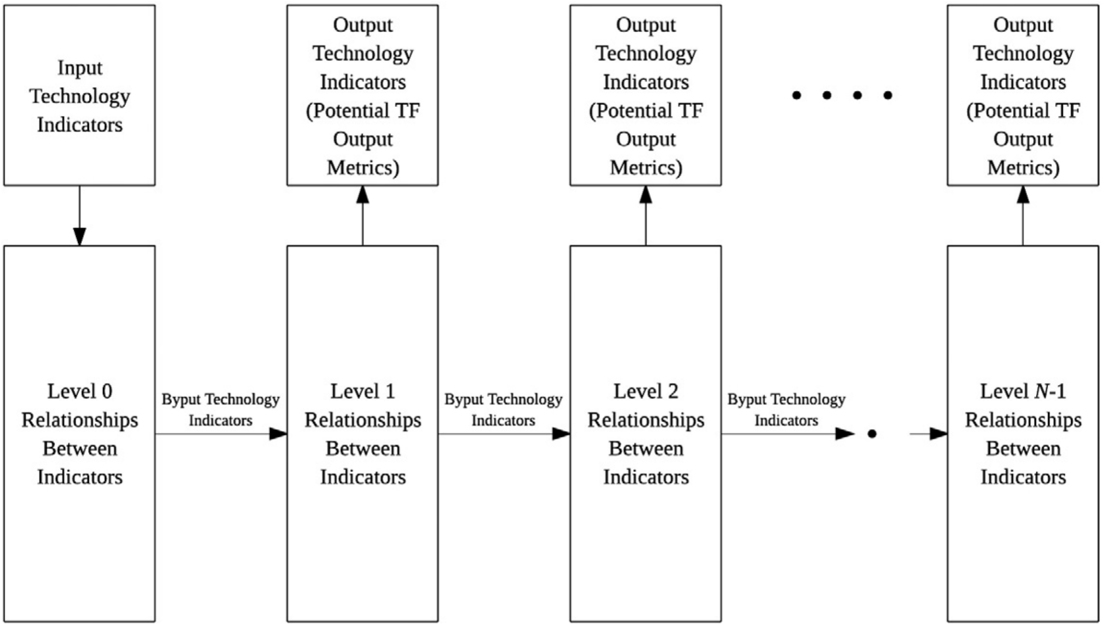

Fig. 3. Relationships between input, byput and output technology indicators.

- *Technology Development Indicators*: This broad grouping includes measures that track the development, change, progress and trend of a technology from a technological perspective (Nyberg and Palmgren, 2011).
- *Market Development Indicators*: This broad grouping includes all indicators related to the market development and potential application areas of the technology, including sales, investment and industrial applications (Nyberg and Palmgren, 2011).

The proposed framework for SEM based DF for TF allows for the use of any of the above stated types of technology indicators as latent or formative indicators for endogenous and exogenous constructs in the model. More specifically, input technology indicators are used with exogenous constructs. Conversely, bypass and output indicators are used for endogenous constructs. The TF output metrics, which will be used by decision makers to drive strategic action, consist of output metrics related to endogenous constructs in the SEM model. External environment related indicators contributing to exogenous constructs that realize context sensitivity in the DF process could also include technology indicators. For example, Sohn and Moon (Sohn and Moon, 2003) used the Technology Commercialization Success Index (TCSI) metric, which is an example of a market prospect indicator in the framework proposed by Watts and Porter (Watts and Porter, 1997), as the primary TF output metric for their SEM model.

2.3. Proposed framework for SEM based DF for TF

Sohn and Moon showed in (Sohn and Moon, 2003) that SEM can be used as an effective regression technique to evaluate a multi-layered hierarchal model through progressive aggregations and refinements of input technology indicator data in order to produce a reliable statistical estimate of the TCSI TF output metric. By extending Soon and Moon's (Sohn and Moon, 2003) use of SEM for TF through Steinberg's use of SEM to implement context sensitive DF (Steinberg, 2009), this paper proposes the framework depicted in Fig. 4 for SEM based context sensitive DF for TF. The framework was developed through inductive reasoning by overlaying Fig. 3′s relationship framework for technology indicators, as defined by Grupp (Grupp, 1998), on the general context sensitive DF framework of Fig. 2 and applying the SEM construct grouping and layering framework of Fig. 1.

In this framework multi-layered aggregation and refinement of technology and context related information is accomplished by the processing performed at DF Levels 0 through *N*-1, where *N* is user selected. The number of levels *N* will be determined not only by the complexity of the technology domain under consideration, but also by time and cost constraints. Also, potential diminishing returns resulting from additional levels of aggregation and refinement will also be determining factors in defining *N*.

Input technology indicators (Nyberg and Palmgren, 2011; Grupp, 1998) and context related indicators (Steinberg, 2009) are used as inputs to technology related endogenous constructs and context related exogenous constructs, respectively. Note that the use of bi-directional interconnections between indicators and constructs, as well as between multiple constructs, is based on the SEM path diagram conventions defined in (Staphorst, 2010). This illustrates that positive or negative correlation can exist between constructs, as well as the fact that indicators can be either reflective or formative in nature.

To gain insight into the functioning of this framework, consider the aggregation and refinement that occur in progressing from DF Level 0 to DF Level 1: Regression analysis outputs generated for the

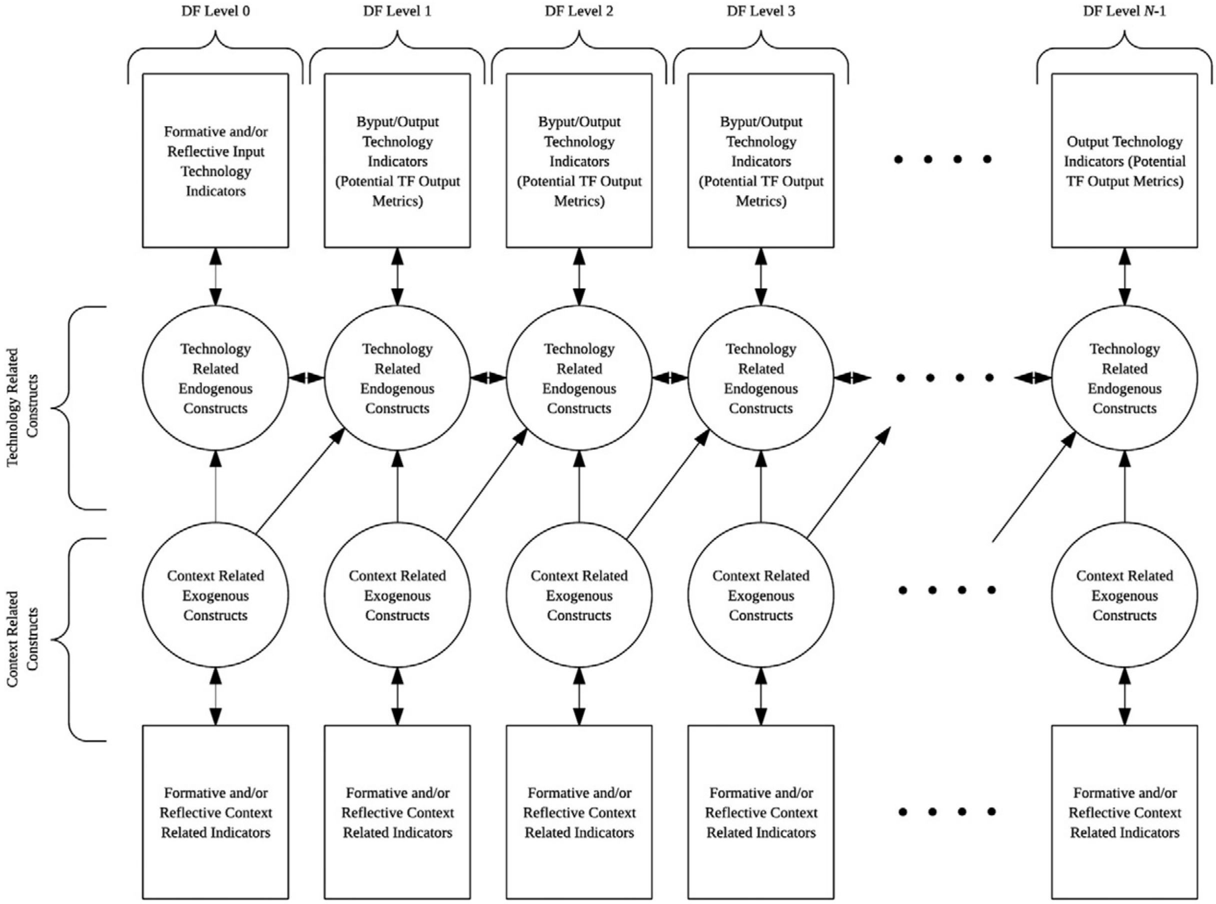

Fig. 4. Proposed Framework for SEM Based DF for TF.

technology related exogenous constructs at DF Level 0 contribute formatively or reflectively to technology related endogenous constructs at DF Level 1. Regression analysis outputs for the context related exogenous constructs of DF Level 0 contribute to context related exogenous and technology related endogenous constructs at DF Level 1. The regression analysis results produced at DF Level 1 for context related exogenous constructs can also contribute to technology related endogenous constructs at this same level. Technology indicators for the technology related constructs at DF Level 1 could potentially be selected as the TF output metrics, or could simply be byput technology (Nyberg and Palmgren, 2011; Grupp, 1998) indicators if additional DF levels are required for further aggregation and refinement. The aggregation and refinement achieved by moving from DF Level *x*-1 to DF Level *x*, for $x = 1, 2, 3, \ldots, N\text{-}1$, follows a similar interconnection structure as the progression from DF Level 0 to DF Level 1, with the exception that now constructs at DF Level *x*-1 contribute to constructs at DF Level *x*.

2.4. Construction and utilization of a model instantiation of the framework

Constructing a model instantiation of Fig. 4′s proposed framework for SEM based DF for TF (referred to as SEM model building in the remainder of this study), involves firstly defining an appropriate set of technology related endogenous constructs and context related exogenous constructs, with relevant technology and context related measurement indicators, respectively, which reflects the fundamental characteristics of the technology domain under investigation. This is then followed by defining a set of hypothesized relationships between these constructs, emanating from theory, action research, colloquial knowledge or speculation. The last stage in the SEM model building exercise involves using empirical data, captured for the each of the technology and context related measurement indicators, in a PLS regression analysis (Vinzi et al., 2010) in order to determine the significance (including indicator loadings) of these measurement indicators, as well as the significance (including the path coefficients) of the defined set of hypothesized relationships.

Exploratory research efforts that make use of SEM frequently terminate at this final stage of SEM model building, as the research objectives for such studies typically involve testing the significance of the hypothesized relationships in the model instantiation (Vinzi et al., 2010). In the case of the proposed framework for SEM based DF for TF, such hypothesis testing could prove especially useful in determining the impact of changing contextual factors, such as technology policy decisions, on the technology related endogenous constructs defined for the technology domain under investigation. As such, the proposed framework effectively provides one with a capability to forecast the relational influence between technology related endogenous constructs and context related exogenous constructs.

A model instantiation of the proposed framework, constructed using the SEM model building process described above, can also be used to forecast TF output metrics through a process best referred to as SEM post-processing (Vinzi et al., 2010). This process involves populating the structural equations defined by the SEM model instantiation with the known context and input/byput technology indicator data from a single metric measurement snapshot and solving these equations to determine the unknown TF output metrics by calculating their associated output technology indicator values.

3. Constructing an example NREN model instantiation of the framework

NREN's are frequently used as incubators for the development of new networking technologies and services (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012). Hence, NREN's contribute significantly to the creation of new Internet based business ventures, innovative business models and game changers in the way society works and plays. For example, Facebook and Google have their roots within the NREN environments of Harvard University and Stanford University, respectively.

TERENA, which now forms part of GÉANT, was an was a not-for-profit association of European NRENs with the objective to provide a platform for NREN's to collaborate and openly share knowledge on networking technologies, services and infrastructure. TERENA performs an extensive yearly survey amongst the global NREN community in order to determine current technology and services trends. The results and interpretation of these surveys are then openly published as part of TERENA's NREN compendium series.

The NREN model instantiation example detailed in the following subsections, constructed using the model building process described in Section 2.4, was created using insights captured in TERENA's NREN compendium for 2012 (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012), knowledge gained through action research (Gustavsen, 2008) performed by the authors during their involvement with the management and operations of SANReN (The South African National Research Network (SANReN), 2013), as well as hypotheses presented and tested in peer-reviewed literature. It is an improved version of the example NREN model instantiation originally proposed and analyzed in (Staphorst et al., 2014). Note that the use of this example NREN model instantiation to performed TF in this study, presented as part of the analysis results discussed in Section 5, was limited to forecasting the relational influence between technology related endogenous constructs and context related exogenous constructs. Forecasting of TF output metrics using SEM post-processing of the example NREN model instantiation will be considered during future research.

3.1. Overview of the example NREN model instantiation

Fig. 5 presents an example NREN technology domain model instantiation of Section 2′s framework employing the SEM path diagram conventions defined in (Staphorst, 2010). This example model instantiation employs $N = 3$ DF levels. Level 0, Level 1 and Level 2 focus on NREN Connectivity (i.e. the NREN provided infrastructure to deliver advanced services), NREN Services (i.e. the portfolio of advanced services provided to users in order to make use of the NREN provided infrastructure) and NREN Utilization (i.e. a measure use of the NREN provided through the advanced services available to users), respectively.

In essence the NREN Connectivity level is an aggregation of Layer 1 (Physical) through to Layer 6 (Presentation layer) in the 7-layered Open Systems Interconnection (OSI) model (Zimmerman, 1980), while the NREN Services level represents Layer 7 (Application layer). The 7-layered OSI model has been unofficially extended through the addition of Layers 8 to 10, representing Human-Computer Interaction (HCI) related aspects (Bauer and Patrick, 2004). NREN Utilization is one possible representation of these HCI related layers.

At Level 0 of the example NREN model instantiation, which focuses on infrastructure related technology metrics, a single technology related endogenous construct, namely *NREN Infrastructure Capability* (η_1), is defined. The purpose of this construct is to model the extent to which the NREN has invested in dark fiber infrastructure and managed circuits (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011; Savory, 2012). Dark fiber is defined fiber infrastructure that is either owned outright by the NREN, or where the NREN has secured a long-term Indefeasible Right of Use (IRU) for the use of fiber (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011). Managed circuits are fiber infrastructure owned by another party and leased by the NREN (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011). Based on (Savory, 2012), in

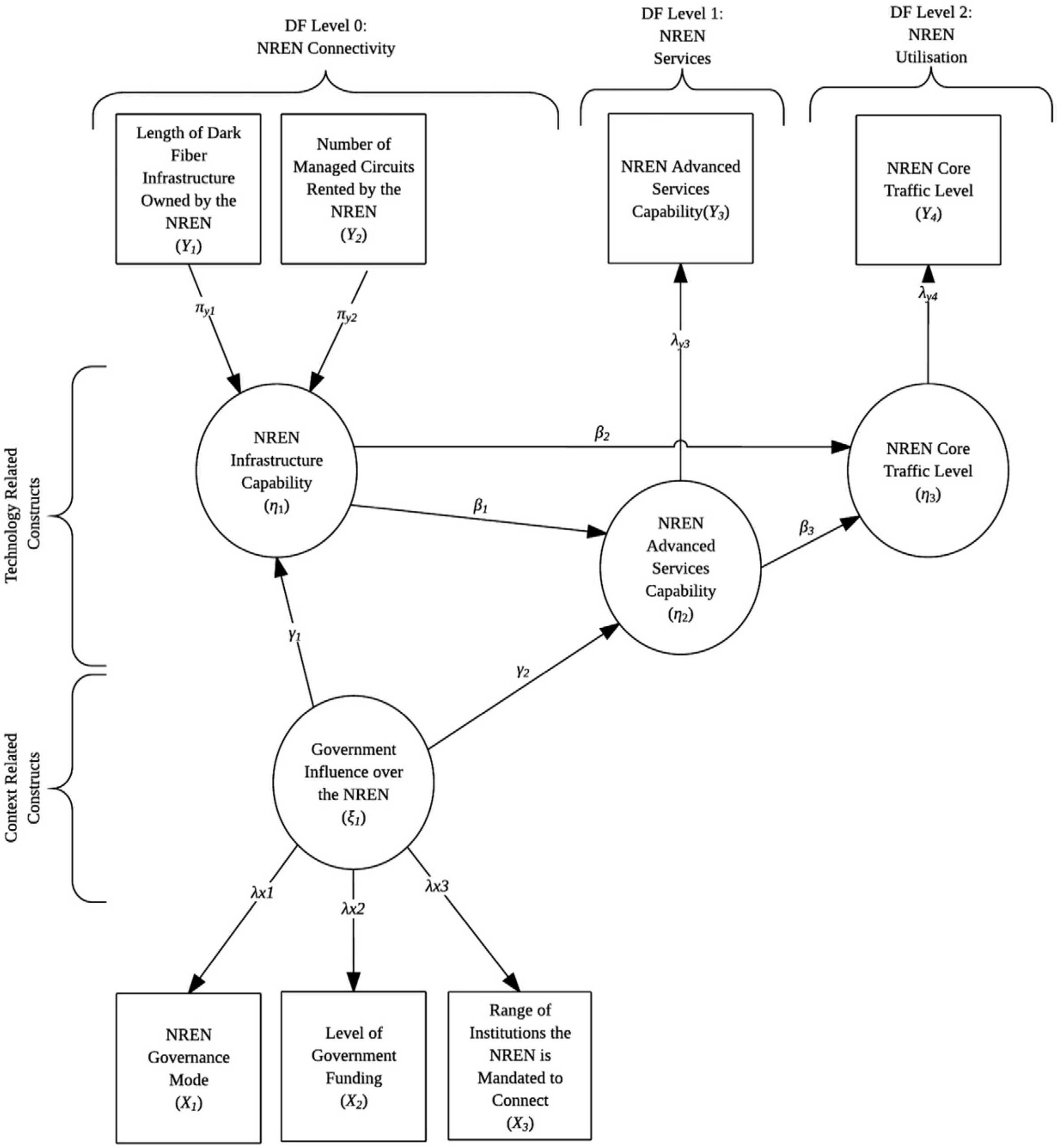

Fig. 5. Example model instantiation for the NREN technology domain.

the example NREN model instantiation it is postulated that the *NREN Infrastructure Capability* (η_1) construct will be related to two formative input technology indicators (i.e. both indicators jointly represent the construct) that measure the length of available dark fiber infrastructure (denoted as *Length of Dark Fiber Infrastructure Owned by the NREN* (Y_1) with indicator loading π_{y1}) and the number of rented managed circuits (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011) (denoted as *Number of Managed Circuits Rented by the NREN* (Y_2) with indicator loading π_{y2}), respectively.

Also defined at Level 0 is a single context related exogenous construct entitled *Government Influence over the NREN* (ξ_1). This construct has three reflective indicators (i.e. each indicator is capable of individually representing the construct) that measure the NREN governance mode (denoted as *NREN Governance Mode* (X_1) with indicator loading λ_{x1}), level of government funding provided to the NREN (denoted as *Level of Government Funding* (X_2) with indicator loading λ_{x2}) and the range of institutions the NREN is mandated to connect (denoted as *Range of Institutions the NREN is Mandated to Connect* (X_3) with indicator loading λ_{x2}), respectively. NREN governance mode can range from full government driven governance through to no government driven governance (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; Staphorst, 2010). The range of institutions that the NREN is mandated to connect can vary from only type of institutions, such as universities, to a suite of various types of institutions, such as research organizations, universities, schools, etc. (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012). A positive relation between *Government Influence over the NREN* (ξ_1) and *NREN Infrastructure Capacity* (η_1) is postulated and represented by the path coefficient γ_1. This relation was derived from the notion that government intervention is required at various points in the NREN value chain, such as infrastructure funding, policy definition, regulation, etc. in order to ensure that an NREN successful matures in terms of the connectivity and advanced services that it provides (Janz and Kutanov, 2012; Greaves, 2009).

It is important to note that additional context related measurement indicators and constructs from the political, economic, sociological, legal and environmental domains can be added to a model instantiation such as this example NREN model instantiation in order to potential improve the model's ability to forecast output technology metrics. However, given that this example NREN model instantiation was tested using the data available from the 2011 TERENA NREN compendium, the context related measurement indicators were limited to those associated with the *Government Influence over the NREN* (ξ_1) construct.

Level 1 of the example NREN model instantiation, which focuses on services related technology metrics, defines a single exogenous technology related construct entitled *NREN Advanced Services Capability* (η_2). This construct embodies the NREN's capability to provide a suite of advanced NREN services (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011), such as Authentication and Authorization Infrastructure (AAI) services, provisioning of Identity Federation Services, hosting of Identity Federation Services and inter-federating with other NRENs (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012). The construct has a single reflective byput technology metric as reflective indicator, measuring the size of the portfolio of advanced services offered and hosted by the NREN (denoted as *NREN Advanced Services Capability* (Y_3) with indicator loading λ_{y3}). A postulated positive relationship between *NREN Infrastructure Capability* (η_1) and *NREN Advanced Services Capability* (η_2) is represented by the path coefficient β_1. This relationship emanates from the postulation in (Greaves, 2009) that an NREN requires an advanced infrastructure capability in order to be able to deliver a portfolio of advanced services.

While no exogenous context related construct is defined for Data Fusion Level 1, it is postulated that the Level 0′s *Government Influence over the NREN* (ξ_1) is positively related to the NREN's ability to deliver advanced services (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012). This relationship is captured in the SEM model of Fig. 5 by means of path coefficients γ_2. The postulated relationship is based on the reasoning in (Janz and Kutanov, 2012; Greaves, 2009) that government intervention is required at various points in the NREN value chain in order to ensure that an NREN successful matures in terms of the advanced services portfolio that it provides.

Level 2 in the example model instantiation focuses on the utilization of the NREN, which is frequently used as a proxy to measure the impact that an NREN creates in its beneficiary communities (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; The South African National Research Network (SANReN), 2013), as well as the Return of Investment (ROI) of the funders of the NREN (Bech, 2011). A single context related exogenous construct, entitled *NREN Core Traffic Level* (η_3), which represents the bandwidth usage in the core network of the NREN, is used to represent the utilization of the NREN (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011). This construct is directly measure by means of the reflective measurement indicator *NREN Core Traffic Level* (Y_4), with indicator loading λ_{y4}. This measurement indicator is also the technology forecasting output metric for the example NREN model instantiation.

Postulated positive relationships between *NREN Infrastructure Capability* (η_1) and *NREN Core Traffic Level* (Y_4), as well as between *NREN Advanced Services Capability* (η_2) and *NREN Core Traffic Level* (Y_4), are represented by path coefficients β_2 and β_3, respectively. The positive relationship between infrastructure capability and network utilization (i.e. core network traffic level) is supported in (Savory, 2012). In (Janz and Kutanov, 2012; Greaves, 2009) it is postulated that the maturity of the advanced service portfolio is a driver in the utilization of broadband networks, thereby justifying the positive relationship between the advanced services capability and network utilization.

3.2. Research propositions for the example NREN model instantiation

The postulated relationships between constructs in Fig. 5′s example NREN model instantiation give rise to the set of research propositions below, which are evaluated in Section 5.3. These research propositions' association with the various paths defined in the example NREN model instantiation is detailed in Fig. 5, as well as Table 3.

Research Proposition H1. The NREN's infrastructure capability is positively related to the level of government influence over the NREN. This hypothesis stems from notion that government influence is required in order to ensure that an NREN is successful in maturing its infrastructure capability (Janz and Kutanov, 2012; Greaves, 2009).

Research Proposition H2. The advanced services capability of the NREN is positively related to the NREN's infrastructure capability. This hypothesis is supported by the postulation in (Greaves, 2009) that an NREN requires an advanced infrastructure capability in order to be able to deliver a portfolio of advanced services.

Research Proposition H3. The advanced services capability of the NREN is positively related to the level of government influence over the NREN. This hypothesis stems from notion that government influence is required in order to ensure that an NREN is successful in maturing its advanced services portfolio (Janz and Kutanov, 2012; Greaves, 2009).

Research Proposition H4. The level of core network traffic in the NREN is positively related to the infrastructure capability of the NREN, as postulated in (Savory, 2012).

Research Proposition H5. The level of core network traffic in the NREN is positively related to the advanced services capability of the NREN, as postulated in (Janz and Kutanov, 2012; Greaves, 2009).

4. Analysis results for the example NREN model instantiation

Secondary data from TERENA's NREN compendiums for 2011 (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011) was used to determine Fig. 5′s indicator loadings and path coefficients through PLS regression analysis. Table 1 below summarizes the composition of the NREN model instantiation indicator data using the secondary data extracted from the 2011 TERENA NREN compendium (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011). A total of 61 NRENs responded to TERENA's survey to collect data for this. The original survey distributed by TERENA to NRENs is available from (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011).

In this study the SmartPLS (Ringle et al., 2005) freeware software package was employed to realize the example NREN model instantiation of Fig. 5 and calculate all loadings and path coefficients through PLS regression. SmartPLS was configured to normalize all indicator data, as a variety of scaling approaches and ranges was used by TERENA in collecting the original data. SmartPLS was also used to evaluate the reliability and validity test criteria defined in (Staphorst et al., 2013) with the results discussed in Section 4.2. Note that only 28 NRENs provided all of the survey inputs in order to calculate the indicator inputs according to Table 1. Hence missing data was flagged and SmartPLS configured to use a mean replacement algorithm to compensate for this (Ringle et al., 2005).

4.1. Measurement indicator and path coefficient results

The reporting of the PLS regression results for the example NREN model instantiation, presented in the following subsections, was based

Table 1
Technology and context related indicator data composition.

Technology or Context Related Indicator	*Indicator composition*
NREN Governance Mode (X_1)	Extracted from the online profiles of the respondent NRENs of the 2011 compendium (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011) using following the scaling: • The NREN is a government agency or part of a ministry = 3 • Government appoints at least half of the NREN's governing body = 2 • Indirect relationship between the NREN and government = 1 • No formal relationship between the NREN and government = 0
Level of Government Funding (X_2)	Level of government funding (as a percentage of total funding) received by respondent NRENs, as summarized in Graphs 6.4.2 and 6.4.3 in (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011)
Range of Institutions the NREN is Mandated to Connect (X_3)	Sum of the institution types supported by respondent NRENs, as shown in Table 2.2.1 of (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011)
Length of Dark Fiber Infrastructure Owned by the NREN (Y_1)	Total length of dark fiber [in kilometers] owned by respondent NRENs as summarized in Table 3.6.3 of (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011)
Number of Managed Circuits Rented by the NREN (Y_2)	Total number of managed circuits rented by respondent NRENs as summarized in Table 3.3.2 of (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011)
NREN Advanced Services Capability (Y_3)	Total number of positive answers to the following questions in Table 5.3.1.1 in (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011): • Does the NREN provide of Authentication and Authorization Infrastructure (AAI) services? • Does the NREN provide Identity Federation services? • Does the NREN operate the Identity Federation services? • Does the NREN's Identity Federation services inter-federate with those provided by other NRENs?
NREN Core Traffic Level (Y_4)	Annual level (measured in terabytes per year) of traffic sent on to the backbone networks of respondent NRENs, as measured by $T_1 + T_4$ in Graphs 4.2.1 and 4.2.2 in (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011)

on the reporting standard defined by Vinzi, Chin, Henseler and Wang (Vinzi et al., 2010). According to this reporting standard, the PLS regression results for the measurement portion of the SEM path diagram, consisting of the loadings for all of the measurement indicators in the model, are reported first, followed by the PLS regression results for the structural portion of the SEM path diagram, consisting of the path coefficients for all interrelationships between constructs.

4.1.1. Measurement portion SEM regression results

The indicator loadings for the measurement portion of the example NREN model instantiation, determined using SmartPLS (Ringle et al., 2005), are listed in Table 2. Although these loadings were not used directly in order to evaluate the research propositions stated in Section 3.2, a detailed investigation thereof was crucial in order to determine those reflective indicators that did not comply with the minimum Indicator Reliability level of 0.4 (see Section 5.2.1).

4.1.2. Structural portion SEM regression results

The path coefficients for the structural portion of the example NREN model instantiation, which were determined using SmartPLS (Ringle et al., 2005), are listed in Table 3. Significance testing for these path coefficients, based on asymptotic *t*-statistics, is presented in Section 5.2.2. These path coefficients and their associated significance test results were used in Section 5.3 to evaluate the research propositions listed in Section 3.2.

4.2. Reliability and validity analysis for the example NREN model instantiation

Similar to the reporting standard for SEM indicator loading and path coefficient results Vinzi, Chin, Henseler and Wang (Vinzi et al., 2010) suggest that the reporting of reliability and validity test results first considers the measurement portion, which include Indicator Reliability, Construct Reliability and Convergent Validity (Staphorst et al., 2013). This is then followed by the structural portion, which includes Coefficients of Determination, Path Coefficient Significance and Predictive Validity (Staphorst et al., 2013). The reasoning behind this approach is that a lack in confidence in the accuracy and representivity of the measurement indicators in a model instantiation negates the need to test the reliability and validity of the structural portion (Vinzi et al., 2010).

4.2.1. Measurement portion reliability and validity results

This subsection details the reliability and validity test results, determined using SmartPLS (Ringle et al., 2005), for the measurement portion of the SEM for the example NREN model instantiation, based on the metrics defined in (Vinzi et al., 2010) and detailed in (Staphorst, 2010). Table 4 presents the Indicator Reliability judgment (Staphorst, 2010), Construct Reliability (Staphorst, 2010) and Convergent Validity (Staphorst, 2010) test results.

The Indicator Reliability test (Staphorst, 2010), which gives an indication of the level of variance in the measurement indicator that can be explained by its associated latent construct (Staphorst, 2010), revealed that none of the reflective indicators exhibited a loadings <0.4 during a first-run PLS regression SEM analysis. As a result, no reflective indicators had to be removed and all subsequent SEM analyses could be performed on the model instantiation as is. All formative indicators were retained regardless of their loadings, since the concept of Indicator Reliability is not applicable to formative indicators. This is because these indicators can exhibit low correlation with their associated latent constructs, but still contribute significantly to their overall variance (Vinzi et al., 2010).

Construct Reliability (Staphorst, 2010), which considers whether the set of reflective indicators associated with a latent construct jointly measure it adequately (Vinzi et al., 2010), employs both the classic Cronbach's Alpha metric (Vinzi et al., 2010) and the more contemporary Composite Reliability measure (Fornell and Larcker, 1981). This study's final judgment on the adequacy of a set of reflective indicators to measure their related latent construct was based on the requirement that the Composite Reliability measure needs to exceed a minimum level of 0.6 (Vinzi et al., 2010). As is clear from Table 4 the only latent construct with reflective indicators present in the model was *Government Influence over the NREN* (ξ_1), which complied with this requirement for Composite Reliability (Fornell and Larcker, 1981).

Convergent Validity considers the correlation between responses obtained by maximally different methods of measuring the same construct (Vinzi et al., 2010). It is determined through the Average Variance Extracted (AVE) metric (Staphorst, 2010; Fornell and Larcker, 1981), which measures the variance of each latent construct's reflective indicators (as captured by the construct itself) relative to the total measured variance. Measured against the study's elected threshold value of 0.5 for this metric (Staphorst, 2010), it can be concluded from Table 4′s results that the reflective indicators of the only latent construct *Government Influence over the NREN* (ξ_1) exhibited a sufficient AVE level, indicating that for this

Table 2
Measurement portion indicator loading results.

Constructs	*Type*	*Measurement indicators*	*Loadings*
Government Influence over the NREN (ξ_1)	Reflective	*NREN Governance Mode* (X_1)	$\lambda_{x1} = 0.892$
	Reflective	*Level of Government Funding* (X_2)	$\lambda_{x2} = 0.805$
	Reflective	*Range of Institutions the NREN is Mandated to Connect* (X_3)	$\lambda_{x3} = 0.854$
NREN Infrastructure Capability (η_1)	Formative	*Length of Dark Fiber Infrastructure Owned by the NREN* (Y_1)	$\pi_{y1} = 0.473$
	Formative	*Number of Managed Circuits Rented by the NREN* (Y_2)	$\pi_{y2} = 0.737$
NREN Advanced Services Capability (η_2)	Reflective	*NREN Advanced Services Capability* (Y_3)	$\lambda_{y3} = 1.0$
NREN Core Traffic Level (η_3)	Reflective	*NREN Core Traffic Level* (Y_5)	$\lambda_{y4} = 1.0$

construct the majority of the total variance measured was due to indicator variance and not due to measurement error.

Discriminant Validity for the measurement portion of a SEM model considers the level of dissimilarity in the measurements obtained by the measurement tool for different constructs (Vinzi et al., 2010). A necessary condition to achieve Discriminant Validity requires that the shared variance between a latent construct and its indicators (determined by taking the square root of it AVE) exceeds the shared variance between this latent construct and any other latent constructs. Since the example NREN model instantiation only had one latent construct with reflective indicators, namely *Government Influence over the NREN* (ξ_1), this test was unnecessary.

4.2.2. Structural portion reliability and validity results

The results for the reliability and validity tests for the structural portion of example NREN model instantiation, based on the metrics defined in (Vinzi et al., 2010) and detailed in (Staphorst, 2010), are presented in this subsection. Table 5 details the Path Coefficient significance test results, while Table 6 considers the Coefficients of Determination and Predictive Validity test results. These results were obtained using SmartPLS (Ringle et al., 2005).

As with covariance based multiple regression techniques, the quality of the structural portion of a model instantiation can be investigated by means of a bootstrapping procedure (Vinzi et al., 2010) in order to determine the significance levels of the path coefficients. The significance of path coefficients (also sometimes referred to as Goodness-of-Fit) can be tested via asymptotic *t*-statistics. From Table 5′s Path Coefficient Significance test results, obtained using SmartPLS's bootstrapping function (Ringle et al., 2005) configured to generate 500 sets of subsamples from the 61 cases in the original sample, the only path that exhibited a *p*-value larger than the maximum acceptable significance level of $\alpha = 0.10$ was *NREN Infrastructure Capability* (η_1) → *NREN Advanced Services Capability* (η_2). Hence, this path was deemed insignificant.

The Coefficients of Determination (R^2) reflect the share of an endogenous construct's variance explained by related endogenous or exogenous constructs (Vinzi et al., 2010). The test results given in Table 6 revealed that the interrelationships between the endogenous latent constructs and their related constructs produced explained variances exceeding the minimum level of 0.1 (Staphorst, 2010). Also, interrelationships with the *NREN Advanced Services Capability* (η_2) endogenous latent construct were deemed to be strong, since the R^2 for this construct exceeded 0.7 (Staphorst, 2010). Interrelationships with the *NREN Core Traffic Level* (η_3) endogenous latent construct were viewed as weak, since the R^2 for this construct was lower than 0.3 (Staphorst, 2010).

Table 3
Structural portion path coefficient results.

Research Proposition: SEM path	*Coeff.*
H1: *Government Influence over the NREN* (ξ_1) → *NREN Infrastructure Capability* (η_1)	$\gamma_1 = 0.599$
H2: *NREN Infrastructure Capability* (η_1) → *NREN Advanced Services Capability* (η_2)	$\beta_1 = 0.016$
H3: *Government Influence over the NREN* (ξ_1) → *NREN Advanced Services Capability* (η_2)	$\gamma_2 = 0.855$
H4: *NREN Infrastructure Capability* (η_1) → *NREN Core Traffic Level* (η_3)	$\beta_2 = 0.289$
H5: *NREN Advanced Services Capability* (η_2) → *NREN Core Traffic Level* (η_3)	$\beta_3 = 0.187$

In order to determine the Predictive Validity of a model instantiation the Stone-Geisser (referred to as Q^2) non-parametric test is performed (Staphorst, 2010; Vinzi et al., 2010) based on a blindfolding procedure (Zikmund, 2003). The model instantiation is considered to have Predictive Validity if $Q^2 > 0$ (Vinzi et al., 2010). The Stone–Geisser test criterion can take on two distinct forms, depending on the type of prediction that is investigated: The first form, which is geared at determining the Predictive Validity of the measurement portion (although usually calculated during the structural portion's validity evaluation), is referred to as the Cross-validated Communality (Vinzi et al., 2010) and is denoted by H^2. Cross-validated Communality measures the ability of the model instantiation to predict the observable endogenous constructs from their own latent construct scores (Vinzi et al., 2010). The second form, which evaluates the Predictive Validity of the structural portion, is referred to as Cross-validated Redundancy (Vinzi et al., 2010). This metric, denoted by F^2, measures the model instantiation's ability to predict the observable endogenous constructs using latent constructs that predict the block of data in question (Vinzi et al., 2010).

A review of the Predicative Validity test results for the *NREN Core Traffic Level* (η_3) construct, directly observable via the output forecasting technology metric of interest *NREN Core Traffic Level* (Y_4), revealed that both the Cross-validated Communality (H^2) and the Cross-validated Redundancy (F^2) tested positively. Hence, the both the example NREN model instantiation's measurement indicators and the defined structural relationships are well suited to forecasting the NREN's core network traffic level.

4.3. Evaluation and discussion of the framework outputs for the example NREN model instantiation

Using the calculated path coefficients in Table 3 and the path coefficient significance test results in Table 5, the research propositions defined for the example NREN model instantiation in Section 3.2 were evaluated as follows:

Research Proposition H1. The path coefficient of $\gamma_1 = 0.599$ supports the direction of the proposed relationship between *Government Influence over the NREN* (ξ_1) and *Infrastructure Capability* (η_1) *Government Influence of the NREN* (ξ_1). Furthermore, the path coefficient was judged to be significant at the maximum allowed significance level of $\alpha = 0.10$. Hence, this hypothesized relationship was not rejected and supports the notion in (Janz and Kutanov, 2012)(Greaves, 2009) that government influence is required in order to ensure that an NREN is successful in maturing its infrastructure capability. Moreover, the positive influence that the government has over the infrastructure capabilities of an NREN was to be expected, since most NRENs are government interventions geared are supporting a country's research and education communities by enhancing the available research and education infrastructure and services (Janz and Kutanov, 2012)(Greaves, 2009).

Research Proposition H2. This hypothesized relationship between *NREN Infrastructure Capability* (η_1) and *NREN Advanced Services*

Table 4
Indicator Reliability, Construct Reliability and Convergent Validity Test Results.

Constructs	*Measurement indicators*	*Indicator Reliability judgment*	*Construct Reliability*		*Convergent Validity*
			Cronbach's Alpha	*Composite Reliability*	
Government Influence over the NREN (ξ_1)	*NREN Governance Mode* (X_1)	Included	0.810	$\rho_{\xi,1} = 0.887$	$AVE_{\xi,1} = 0.724$
	Level of Government Funding (X_2)	Included			
	Range of Institutions the NREN is Mandated to Connect (X_3)	Included			
NREN Infrastructure Capability (η_1)	*Length of Dark Fiber Infrastructure Owned by the NREN* (Y_1)	Included	Tests not applicable: This construct has formative indicators (Staphorst, 2010)		
	Number of Managed Circuits Rented by the NREN (Y_2)	Included			
NREN Advanced Services Capability (η_2)	*NREN Advanced Services Capability* (Y_3)	Included	Tests not applicable: This construct is directly observable (Staphorst, 2010)		
NREN Core Traffic Level (η_3)	*NREN Core Traffic Level* (Y_4)	Included	Tests not applicable: This construct is directly observable (Staphorst, 2010)		

Capability (η_2) was rejected. While the path coefficient $\beta_1 = 0.016$ supported the direction of the proposed relationship, the path coefficient was judged to be not significant at the maximum allowed significance level of $\alpha = 0.10$. This finding is counter to the assertions in (Greaves, 2009) that an NREN requires an advanced infrastructure capability in order to mature its portfolio of advanced services. This can be explained by noting that, in the telecommunications industry providers in developing countries frequently leapfrog their more developed counterparts by offering advanced services, even though their infrastructure capability might still be nascent (Mbarika et al., 2000). In the case of the NREN community, European NRENs frequently support fledgling NRENs in Africa and Asia to rapidly deploy advanced services through development programs driven by the GÉANT Association (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012)(TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011).

Research Proposition H3. The path coefficient $\gamma_2 = 0.855$ supported the direction of the hypothesized relationship between Government Influence over the NREN (ξ_1) and NREN Advanced Services Capability (η_2). Furthermore, the path coefficient was deemed significant. Hence, this research proposition was not rejected and supports the notion in ((Janz and Kutanov, 2012); (Greaves, 2009)) that government influence is required in order to ensure that an NREN is successful in maturing its advanced services portfolio. As with Research Proposition 1, this finding was to be expected, since most NRENs are government interventions geared are supporting a country's research and education communities by enhancing the available research and education infrastructure and services ((Janz and Kutanov, 2012); (Greaves, 2009)).

Research Proposition H4. The postulated relationship between *NREN Infrastructure Capability* (η_1) and *NREN Core Traffic Level* (η_3) was not rejected since the path coefficient $\beta_2 = 0.289$ was judged to be significant at the maximum allowed significance level of $\alpha = 0.10$. Therefore, the postulated relationship in (Savory, 2012) that an NREN's infrastructure capability is positively related to its usage is supported. This finding correlates with the notion that enhancing NREN infrastructure will lead to improved usage ((TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012); (Greaves, 2009)).

Research Proposition H5. : This research proposition was not rejected, since the path coefficient $\beta_3 = 0.187$ support the direction of the hypothesized relationship between *NREN Advanced Services Capability* (η_2) and *NREN Core Traffic Level* (η_3). Also, this path was deemed significant at the maximum allowed significance level of $\alpha = 0.10$, thereby supporting the notion in ((Janz and Kutanov, 2012); (Greaves, 2009)) that an NREN's advanced services capability is positively related to the usage of the NREN. This finding correlates with the notion that, by providing the beneficiaries of an NREN with a portfolio of advanced services to fully utilize the infrastructure available, the usage of an NREN will increase ((TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012); (Greaves, 2009)).

In summary, the data available from the 2011 TERENA NREN compendium supported all of the postulated relationships in the example NREN model instantiation, with the exception of the postulation that an NREN's advanced services capability is related to its infrastructure capability. This latter result emanated from technology leapfrogging that is prevalent in the NREN community. Furthermore, both the selected input and byput selected metrics, as well as the context related metrics, contributed successfully to the forecasting of the output technology metric measuring the usage of the NREN. The same was also true of the structural relationships defined throughout the three DF levels of the example NREN model instantiation.

5. Conclusions

Applying inductive reasoning to the work of Sohn and Moon (Sohn and Moon, 2003), as well as Steinberg (Steinberg, 2009; Steinberg and Rogova, 2008), the paper derived a framework for SEM based DF for TF. Unlike most TF approaches, the proposed framework not only caters for complex and hierarchical structural relationships between technology indicators and TF output metrics, but also allows for non-linear and non-Gaussian factors and cyclical dependencies amongst model variables, which can be either latent or directly observable (Sohn and Moon, 2003).

An example model instantiation of the proposed framework was presented for the NREN technology domain. This example model instantiation was created using knowledge gained through action

Table 5
Path coefficient significance test results.

Research Proposition: SEM path	*Asymptotic t-Statistic*	*Calculated p-value*	*Significance Judgment for* $\alpha = 0.10$
H1: *Government Influence over the NREN* (ξ_1) → *NREN Infrastructure Capability* (η_1)	3.952	<0.001	Yes
H2: *NREN Infrastructure Capability* (η_1) → *NREN Advanced Services Capability* (η_2)	0.124	0.901	No
H3: *Government Influence over the NREN* (ξ_1) → *NREN Advanced Services Capability* (η_2)	4.840	<0.001	Yes
H4: *NREN Infrastructure Capability* (η_1) → *NREN Core Traffic Level* (η_3)	2.015	0.044	Yes
H5: *NREN Advanced Services Capability* (η_1) → *NREN Core Traffic Level* (η_3)	1.662	0.097	Yes

Table 6
Coefficient of determination and predictive validity test results.

Technology or Context Related Indicator	*Coefficients of Determination* (R^2)	*Predictive Validity* (Q^2)	
		Cross-validated Communality (H^2)	*Cross-validated Redundancy* (F^2)
Government Influence of the NREN (ξ_1)	Test not applicable: Exogenous variable	0.440	Test not applicable: Exogenous variable
NREN Infrastructure Capability (η_1)	0.359	0.054	0.191
NREN Advanced Services Capability (η_2)	0.749	1.0	0.727
NREN Core Traffic Level (η_3)	0.175	1.0	0.142

research in SANReN (Gustavsen, 2008), data captured by TERENA in its yearly NREN compendium series (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012; TERENA COMPENDIUM of National Research and Education Networks in Europe, 2011), as well as hypotheses presented in peer-reviewed literature. The example model instantiation, which consisted of 3 DF levels, suggested that an NREN's infrastructure capability (defined at DF level 0) and advanced services capability (defined at DF level 1) were both positively related to the government influence over the NREN, which was regarded as a contextual influence. Furthermore, the model postulated that an NREN's infrastructure capability was positively related to its advanced services capability and that both the infrastructure and services capabilities were positively related to the usage of the NREN (defined at DF level 2), which was chosen as the technology forecasting output metric of the example model instantiation.

Data from the 2011 TERENA compendium was also used to perform a PLS regression in order to determine the path coefficients and indicator loadings in the example NREN model instantiation. From the PLS regression results obtained for the example NREN model instantiation it can be concluded that all of the selected technology indicators were able to adequately measure their respective technology related model constructs. Furthermore, the TERENA data substantiated all hypothesized relationships in the example NREN model instantiation, except the postulation that an NREN's infrastructure capability is positively related to its advanced services capability. This last finding can be explained by the high prevalence of technology leapfrogging (Mbarika et al., 2000) in the global NREN community (TERENA COMPENDIUM of National Research and Education Networks in Europe, 2012), since developing NRENs are rapidly implementing extensive advanced services portfolios (with the assistance of their more advanced NREN peers), while their infrastructure capability is still somewhat nascent. Lastly, a reliability and validity evaluation of the example NREN model instantiation using the 2011 TERENA compendium data highlighted that both the measurement and the structural portions of the model were capable of contributing adequately to the forecasting of the technology forecasting output metric that measured the usage of an NREN.

Future research activities that will be undertaken include an evaluation of the strengths and weaknesses of the proposed framework for SEM based DF for TF, comparing the proposed framework to various popular TF techniques currently receiving attention from the technology management research community, as well as improving the model instantiation for the NREN technology domain. Strengths that will be explored include the ability of the proposed framework for SEM based DF for TF to utilize contextual information in order to improve its forecasting capability. Weaknesses that will be evaluated include the potential negative impact that poorly defined structural configurations in SEM models have on their resultant Goodness-of-Fit (Vinzi et al., 2010).

Improving the model for the NREN technology domain will entail a two-phase process, with the first phase involving a qualitative study (Zikmund, 2003) that will attempt to identify improved endogenous and exogenous model constructs, technology indicators, as well as interactions between the various indicators and constructs. The unit of analysis (Zikmund, 2003) for this envisioned qualitative phase will be a single NREN, while the population will be all NRENs in existence worldwide at the time of the study. Data collection will be accomplished through the Delphi method (Zimmerman, 1980) using a panel of experts comprising the senior technical managers at leading NRENs from the global community. Analysis of the qualitative data collected through various rounds of engagements with the panel of experts will start with narrative inquiry by means of a process of theme extraction (Staphorst, 2010). This will then be followed by performing a frequency analysis on the extracted themes in order to produce a final set of importance ranked indicators, constructs and interconnections from which the improved NREN model instantiation will be constructed (Staphorst, 2010). Testing the reliability and validity of the collected qualitative data will be accomplished by means of theory triangulation (Staphorst, 2010), as well as data triangulation (Staphorst, 2010) using as baseline published technology indicators from secondary data sources, such as TERENA's NREN compendium series.

The second phase will be quantitative in nature and will aim to determine, using PLS regression analysis, the indicator loadings and path coefficients of the NREN model constructed during the qualitative first phase. As with the qualitative phase the population will be all NRENs in existence at that point in time, with the unit of analysis being a single NREN (Zikmund, 2003). While the data available from the TERENA NREN compendium series will be used as far as possible to populated technology and context related measurement indicators, any additional qualitative data required will be obtained using an online survey consisting of close-ended questions with Likert scaling (Zikmund, 2003), targeted a sample of senior technical managers at the NRENs in the population, selected through a process of convenience sampling (Zikmund, 2003).

Acknowledgements

The Council for Scientific and Industrial Research, as well as the University of Pretoria supported this study through funding contributions. Dr. Pieter de Villiers at the University of Pretoria is recognized for his contribution towards the development the proposed framework for SEM based DF for TF by assisting the authors in the effort apply SEM as a context sensitive DF analysis technique.

References

Bauer, B., Patrick, A.S., 2004. A human factors extension to the seven-layer OSI reference model. http://www.andrewpatrick.ca/OSI/10layer.html (accessed 18/01/2016).

Bech, M., 2011. Calculating the Return on Investment of an NREN. TERENA Networking Conference 2011.

Buchroithner, M., 1998. Geodata Interrelations: Inventory and Structuring Attempt of Taxonomic Diversity. Proceedings of the 2nd Conference on Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images, pp. 11–15.

Chang, H., 2007. A Methodology for the Identification of Technology Indicators, HNI-Verlagsschriftenreihe, 233.

Chang, H., Gausemeier, J., Ihmels, S., Wenzelmann, C., 2008. Innovative Technology Management System with Bibliometrics in the Context of Technology Intelligence. Trends in Intelligent Systems and Computer Engineering: Lecture Notes in Electrical Engineering 6 pp. 349–361.

Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement errors. J. Mark. Res. 18, 39–50.

Greaves, D., 2009. An NREN capability maturity model. http://www.ubuntunet.net/sites/default/files/NREN_Capability_Maturity.pdf (accessed 18/01/2016).

Grupp, H., 1998. Foundations of the Economics of Innovation: Theory, Measurement and Practice. Edward Elgar, Cheltenham.

Gustavsen, B., 2008. Action research, practical challenges and the formation of theory. Action Res. 6 (4), 421–437.

Haenlein, M., Kaplan, A.M., 2004. A beginner's guide to partial least squares analysis. Underst. Stat. 3 (4), 283–297.

Janz, R., Kutanov, A., 2012. The case of NRENs in Central Asia. Cent. Asian J. Glob. Health 1 (1).
Jöreskog, K.G., 1973. In: Goldberger, A.S., Duncan, O.D. (Eds.), A general method for estimating a linear structural equation system, in structural equation models in the social sciences. Seminar, New York.
Lichtenthaler, E., 2004. Technological change and the technology intelligence process: a case study. J. Eng. Technol. Manag. 21, 331–348.
Mack, C.A., 2011. Fifty years of Moore's law. IEEE Trans. Semicond. Manuf. 24 (2), 202–207.
Mbarika, V.W., Byrd, T.A., Raymond, J.E., McMullen, P., 2000. Investments in telecommunications infrastructure are not the panacea for least developed countries leapfrogging growth of teledensity. Int. *J. media* Manag. 2 (3/4), 133–142.
Metcalfe, B., 1995. Metcalfe's law: a network becomes more valuable as it reaches more users. InfoWorld 17 (40), 53–54.
Nielsen, J., 1988. Nielsen's law of internet bandwidth. http://www.nngroup.com/articles/nielsens-law-of-internet-bandwidth/ (accessed 08/01/2013), (April).
Nyberg, A., Palmgren, S., 2011. Using Indicators for Technology Monitoring: Steps Toward a Proposed Framework, Chalmers University of Technology (Master of Science Thesis in the Master Degree Programme: Business Design) .
Porter, A.L., 2005. QTIP: quick technology intelligence processes. Technol. Forecast. Soc. Chang. 72, 1070–1081.
Porter, A.L., 2007. Tech Mining to Drive Open Innovation, Proceedings of the 1st International Conference on Technology Innovation, Risk Management and Supply Chain Management. pp. 1–13.
Porter, A.L., Cunningham, S.W., 2004. Tech Mining: Exploiting New Technologies for Competitive Advantage, Wiley-Interscience 29.
Ringle, C.M., Wende, S., Will, S., 2005. SmartPLS 2.0 (M3) beta. http://www.smartpls.de (accessed 24/01/2014).
Savory, B., 2012. Build vs. Buy: The NREN Network Connectivity Dilemma: About Building Research & Education Fibre Networks. Proceedings and Report of the 5th UbuntuNet Alliance Annual Conference, pp. 160–173.
Sohn, S.Y., Moon, T.H., 2003. Structural equation model for predicting technology commercialization success index (TCSI). Technol. Forecast. Soc. Chang. 70 (9), 885–899.
Staphorst, L., 2010. The Impact of Intellectual Property Rights from Publicly Financed Research and Development on Governance Mode Decisions for Research Alliances. Gordon Institute of Business Science (Master of Business Administration Thesis).
Staphorst, L.Pretorius, Pretorius, L., 2013. Structural Equation Modeling Based Data Fusion for Technology Forecasting: a Generic Framework. Proceedings of PICMET2013 (July).
Staphorst, L.Pretorius, Pretorius, L., 2014. Structural Equation Modeling Based Data Fusion for Technology Forecasting: a National Research and Education Network Example. Proceedings of PICMET2014 (July).
Steinberg, A.N., 2009. Context-sensitive data fusion using structural equation modeling. Proceedings of the 12th IEEE International Conference on Information Fusion (July).
Steinberg, A.N., Rogova, G., 2008. Situation and Context in Data Fusion and Natural Language Understanding. Proceedings of the 11th IEEE International Conference on Information.
TERENA COMPENDIUM of National Research and Education Networks in Europe: 2011 Edition, https://www.terena.org/publications/files/TERENA%20Compendium%202011.pdf (accessed 30/12/2013), 2011.
TERENA COMPENDIUM of National Research and Education Networks in Europe: 2012 Edition, https://www.terena.org/publications/files/TERENA%20Compendium%202012.pdf (accessed 30/12/2013), 2012.
The South African National Research Network (SANReN), 2013. http://www.sanren.ac.za (accessed 27/12/2013).
Vinzi, V.E., Chin, W.W., Henseler, J., Wang, H., 2010. Handbook of Partial Least Squares: Concepts, Methods and Applications. Springer, Berlin, Heidelberg.
Wald, L., 1997. An Overview of Concepts in Fusion of Earth Data. Proceedings of the EARSeL Symposium 1997, Future Trends in Remote Sensing, pp. 385–390.
Wald, L., 1999. Some terms of reference in data fusion. IEEE Trans. Geosci. Remote Sens. 37 (3) (May).
Watts, R.J., Porter, A.L., 1997. Innovation forecasting. Technol. Forecast. Soc. Chang. 56 (1), 25–47.
Wold, H., 1975. Path models with latent variables: the NIPALS approach, in quantitative sociology. In: Blalock, H.M., Aganbegian, A., Borodkin, F.M., Boudon, R., Capecchi, V. (Eds.), International Perspectives on Mathematical and Statistical Modeling. Academic Press, New York, pp. 307–357.
Zikmund, W., 2003. Business Research Methods, 7th Edition. South-Western Cengage Learning, Ohio.
Zimmerman, H., 1980. OSI reference model - the ISO model of architecture for open systems interconnection. IEEE Trans. Commun. 28 (4), 425–432.

Leon Staphorst has been the Manager of the South African Research Network at the Council for Scientific and Industrial Research since January 2013. He started his career in January 2000 as a Lecturer at the University of Pretoria in the areas of systems analysis, digital signal processing, telecommunications, information theory and channel coding. In January 2007 Leon joined Motorola South Africa, initially as a Technical Trainer and thereafter as Training Manager of the Motorola Cellular Training Institute. From June 2008 he was employed as a Customer Solutions Manager at Nokia Siemens Networks South Africa. Leon joined the Council for Scientific and Industrial Research in November 2009, initially as a Principle Research in the area of Radar technology, thereafter as the Group Leader for the Council for Scientific and Industrial Research's Experimental Electronic Warfare Systems group. He holds Bachelors and Masters Degrees in Electronic Engineering (1997 and 2005, respectively), as well as a Masters Degree in Business Administration (2010), all from the University of Pretoria. He is currently studying towards a PhD degree in Technology Management at the University of Pretoria's Graduate School of Technology Management. Leon is registered as a Professional Engineer with the Engineering Council of South Africa, is a Chartered Fellow of the Chartered Management Institute in the United Kingdom, a Fellow Member of the South African Institute for Electrical Engineering and a Senior Member of the Institute of Electrical and Electronic Engineers. He has co-authored 15 technical conference papers and 3 peer reviewed journal papers.

Leon Pretorius obtained the D Eng degree (1983) from the University of Pretoria, South Africa. He has more than 36 years professional, engineering, academic and academic management experience. He was professor at the University of Johannesburg, South Africa until 2007. He is currently Professor in the Graduate School of Technology Management at University of Pretoria. He has concurrently been active as specialist consultant and researcher in engineering industry since 1980. He has supervised more than 170 master and PhD students in engineering as well as engineering and technology management. This includes the supervision of more than 130 successful Master's dissertations and 40 completed Doctoral theses. He is a rated researcher at the National Research Foundation in South Africa. He has also published more than 174 technical conference papers and peer reviewed journal articles as author and co-author in his fields of expertise. He is registered as a Professional Engineer with the Engineering Council of South Africa is an Honorary Fellow of the South African Institute of Mechanical Engineering, Member of the South African Institute for Industrial Engineering, Member of the American Socioty of Mechanical Engineers and Member of the Institute of Electrical and Electronic Engineers.

Tinus Pretorius obtained the M Eng (Engineering Management) and D Eng degrees from the University of Pretoria, South Africa. He has more than 35 years professional, engineering, academic and management experience. After spending 19 years in the energy, defence and automotive industries he joined the university on a full-time basis. He is Professor in Technology Management and also Chairperson of the Graduate School of Technology Management at the University of Pretoria. He is a rated researcher at the National Research Foundation and publishes regularly in international journals and conference proceedings. He supervises PhD and Masters' research students on a regular basis and does consulting work in the field of Technology Management. He is a registered Professional Engineer with the Engineering Council of South Africa and also a Member of the Institute of Electrical and Electronic Engineers.