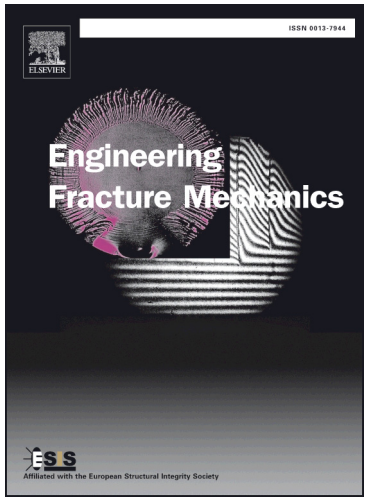
Non-stationary demand forecasting by cross-sectional aggregation

Rostami-Tabar, B, Babai, MZ, Ducq, Y & Syntetos, A

Original citation & hyperlink:
Rostami-Tabar, B, Babai, MZ, Ducq, Y & Syntetos, A 2015, 'Non-stationary demand forecasting by cross-sectional aggregation' International Journal of Production Economics, vol 170, no. A, pp. 297-309.
https://dx.doi.org/10.1016/j.ijpe.2015.10.001

DOI 10.1016/j.ijpe.2015.10.001
ISSN 0925-5273

Publisher: Elsevier



An adaptive stochastic multi-scale method for cohesive fracture modelling of quasi-brittle heterogeneous materials under uniaxial tension

R.M. Sencu, Z. Yang, Y.C. Wang

PII: S0013-7944(16)30058-3
DOI: http://dx.doi.org/10.1016/j.engfracmech.2016.02.040
Reference: EFM 5081

To appear in: *Engineering Fracture Mechanics*

Received Date: 27 September 2015
Revised Date: 17 February 2016
Accepted Date: 18 February 2016

Please cite this article as: Sencu, R.M., Yang, Z., Wang, Y.C., An adaptive stochastic multi-scale method for cohesive fracture modelling of quasi-brittle heterogeneous materials under uniaxial tension, *Engineering Fracture Mechanics* (2016), doi: http://dx.doi.org/10.1016/j.engfracmech.2016.02.040

An adaptive stochastic multi-scale method for cohesive fracture modelling of quasi-brittle heterogeneous materials under uniaxial tension

R. M. Sencu[1], Z. Yang[2*], Y. C. Wang[1]

[1] School of Mechanical, Aerospace and Civil Engineering, University of Manchester, M13 9PL, UK

[2] Faculty of Engineering, Computing and Environment, Coventry University, CV1 2JH, UK

Abstract

An adaptive stochastic multi-scale method is developed for cohesive fracture modelling of quasi-brittle heterogeneous materials under uniaxial tension. In this method, a macro-domain is first discretised into a number of non-overlapping meso-scale elements (MeEs) each of which containing detailed micro-scale finite element meshes. Potential discrete cracks in the MeEs are modelled by pre-inserted cohesive interface elements (CIEs). Nonlinear simulations are conducted for the MeEs to obtain the crack patterns under different boundary conditions. The macro-domain with the same number of overlapped, adaptively size-increasing MeEs are then simulated, until the potential cracks seamlessly cross the boundaries of adjacent MeEs. The resultant cracks, after being filtered by a new Bayesian inference algorithm to remove spurious cracks wherever necessary, are then integrated as CIEs into a final anisotropic macro-model for global mechanical responses. A two-dimensional example of carbon fibre reinforced polymers was modelled under two types of uniaxial tension boundaries. The developed method predicted crack patterns and load-displacement curves in excellent agreement with those from a full micro-scale simulation, but consuming considerably less computation time of the latter.

Keywords: *multi-scale stochastic fracture mechanics; scale coupling; cohesive crack model; overlapping elements; fibre reinforced plastics.*

1. INTRODUCTION

Due to the random distribution of multiple phases from nano-, micro-, meso- to macro-scales, multiphase quasi-brittle materials, such as concrete and fibre-reinforced polymers (FRP) have intrinsically stochastic, heterogeneous and nonlinear physical and mechanical properties across the multi-length scales. As the finer-scale properties directly determine the performance and reliability of structures and systems at coarser-scales, better understanding of these properties and their inter-scale relationships or scale-transferability by multi-scale computational modelling, has become a critical problem (de Borst, 2008; Kassner et al., 2005; Oden et al., 2003; Kanouté et al., 2009; Nguyen et al., 2012a). This is particularly true for fracture problems, as fracture always starts from micro-cracks at strain localization sites, which then propagate, widen and coalesce into meso-cracks and finally discrete macro-cracks. This phenomenon spans a few length scales, demanding a multi-scale modelling approach.

In general, there exist two categories of multi-scale modelling approaches (Belytschko, 2007): (1) concurrent ones, where multiple scales of computing are performed simultaneously; (2) hierarchical or sequential ones, where fine-scale problems are solved in sub-domains and the homogenised results are up-scaled separately in coarse-scale computing.

In modelling fracture problems, the concurrent approaches naturally discretise the regions with strain localisation by fine meshes and the other by coarse meshes to save computational cost. It is particularly useful for problems with a few cracks or weak interfaces known a priori, where the crack-tip and the interfacial regions are modelled in detail for accurate understanding of the fracture mechanism at fine scales (e.g., (Ghosh and Paquet, 2013; Ghosh et al., 2007; Canal et al., 2012; González and Llorca, 2006; Trias et al., 2006a; Trias et al., 2006b; Trias et al., 2006c; Li et al., 2013)). However, for problems with many distributed cracks or unknown cracks, very dense meshes may have to be used in the whole domain to simulate potential cracks, making the concurrent approaches computationally costly.

In the hierarchical approaches, a representative volume element (RVE) or unit cell, based on the classical homogenization theory (Hill, 1963; Hashin, 1965), is assumed to exist at medium scales. Once the existence and size of the RVE is determined by detailed numerical analyses in fine scales, the domain at coarse scales is assumed homogeneous and modelled by a number of RVEs. In doing so, a full analysis of the domain with fine-scale details is avoided. However, recent studies (Phu Nguyen et al., 2010; Gitman et al., 2008; Gitman et al., 2007) find that the RVE exists only in linear-elastic and hardening regimes; once softening occurs as in fracture and damage, the material loses the "representative" properties and the RVE cannot be found, because the material in softening shows localization leading to the loss of statistical homogeneity. If the RVE does not exist, special measures must be taken for multi-scale modelling to maintain the objectivity with respect to the size of the sample cells. Various multi-scale models have been recently developed to solve this dilemma, e.g., the coupled-volume multi-scale model (Gitman et al., 2008), the multi-scale aggregating discontinuities model (Belytschko et al., 2008; Loehnert and Belytschko, 2007), the multi-grid method (Miehe and Bayreuther, 2007; Kaczmarczyk et al., 2010), the homogenization-localization methods (Bosco et al., 2015; Coenen et al., 2012), the enhanced continuous-discontinuous model (Nguyen et al., 2012b; Nguyen et al., 2011), the reduced integration order model (Fish, 2011; Fish

and Shek, 1999), the two-scale homogenization model (Greco et al., 2013; Cusatis and Cedolin, 2007; Desmorat and Lemaitre, 2001), the multi-fractal approach (Carpinteri et al., 2002; Carpinteri and Chiaia, 1997; Xu et al., 2013), and the variational and localized Lagrange multiplier method (Hautefeuille et al., 2012; Markovic and Ibrahimbegovic, 2004).

It can be noted that most of the existing multi-scale models, either concurrent or hierarchical, have tackled fracture problems with single, a few cracks or pre-defined bi-material interfaces. The fine-scale multiphase structures are mostly assumed and the numerical results are difficult to be accurately validated. The latest advances in high-resolution image-based models (Dirrenberger et al., 2014; Ren et al., 2015; Huang et al., 2015) appear very promising for direct validation of the numerical models but they are so far not used in multi-scale modelling and validation. In addition, multi-scale models considering the effects and inter-scale transfer of stochastic information are still largely limited to the prediction of homogenised elastic properties (Xu and Graham-Brady, 2005; Xu and Chen, 2009) rather than complicated fracture evolution. 3D multi-scale modelling of fracture in composite materials has rarely been reported, probably due to the very large number of degrees of freedom in the fine-scale models that are beyond the power of conventional computers. Therefore, much research is still needed to develop more robust multi-scale methods for complicated fracture modelling, as also pointed out by a relatively recent report of USA National Committee on Theoretical and Applied Mechanics (Belytschko, 2007).

As an effort towards overcoming the above deficiencies, a new efficient multi-scale method is developed in this study for modelling complicated fracture behaviour in heterogeneous quasi-brittle materials. In this method, a macro-domain is first discretised into a number of meso-scale elements (MeEs) with detailed micro-scale finite element meshes generated from images obtained by a microscope. The initiation and propagation of multiple cracks is modelled discretely by pre-inserted cohesive interface elements (CIEs) with tension and shear softening laws (Yang et al., 2009; Su et al., 2010). The number and the size of the MeEs required are determined by a size effect study with respective to crack paths and strength. A nonlinear simulation is then conducted separately for each MeE in parallel to obtain the crack patterns under two types of uniaxial tensile boundary conditions. The same number of overlapped, adaptively size-increasing MeEs are then simulated, until the cracks seamlessly cross the boundaries of adjacent MeEs. This can be achieved because the shared microstructures within the neighbouring overlapped elements minimise the potential crack bias and the boundary deformation incompatibility across boundaries. After being filtered by the Bayesian inference algorithm to remove spurious cracks wherever necessary, the resultant cracks are integrated as CIEs (with softening laws mapped from the adaptive meso-scale study) into a final anisotropic macro-model to compute the global responses. The crack patterns and load-displacement curves computed from the developed method are compared with those from a full micro-scale simulation as a means of numerical validation.

It should be noted that the overlapping strategy used herein is similar to the oversampling strategy used in the so-called generalised multi-scale finite element method (GMsFEM) for flow problems in heterogeneous porous media (Hou and Wu, 1997; Chen et al., 2003; Aarnes et al., 2006; Efendiev et al., 2004; Efendiev et al., 2014; Calo et al., 2014). In the GMsFEM, the scale-coupling is realised using multi-scale basis functions containing the deformation information of fine-scale elements. The oversampling strategy uses larger regions than the fine-scale elements to construct more accurate

local basis functions, making the GMsFEM converge faster with fewer degrees of freedom. Although the GMsFEM with oversampling has recently been applied to linear elastic stress analyses (Chung et al., 2014) and seismic wave propagation (Gao et al., 2014), its extension to complicated nonlinear fracture problems is not yet reported. Such an extension is very challenging due to the strain localisation and associated nonlinear material softening, which may presumably make it impossible to form the rigorous analytical derivation of the GMsFEM to ensure deformation compatibility between the fine-scale elements for general external boundary conditions.

As the first attempt to use the overlapping strategy for multi-scale cohesive fracture modelling in quasi-brittle materials, this study will investigate problems with randomly and evenly distributed inclusions and predominantly mode-I fracture, so that all the meso-scale elements are assumed under uniaxial tension and can be modelled independently. The overlapping grids with common areas are used to enhance the deformation compatibility. The scale transfer is realised by the energy conservation principle through mapping the traction-displacement softening curves in CIEs from the MeEs to the MaEs. The effects of inclusion-matrix interfacial fracture properties and two types of boundary conditions on the crack path and strength are also investigated. A two-dimensional example of FRP plates is modelled as a case study.

2. METHODOLOGY

2.1. Framework

Figure 1 shows the framework of the developed method with the key modules highlighted. It starts with acquiring a high-resolution image of the global domain by micro-tests using advanced techniques such as high-resolution cameras, microscopes and X-ray Computer Tomography (XCT). The image is then processed and segmented into different phases. If the crack paths and load-carrying capacities are not available from the micro-tests, a full micro-scale FE modelling of the global domain is carried out to validate the multi-scale modelling. In this case, the global-domain image is first transformed into FE meshes of solid elements. Cohesive interface elements with softening traction-displacement constitutive laws in normal and shear directions are then inserted into the matrix mesh and between the matrix-inclusion interfaces, to model potential cracks. This method has proved to be very effective in modelling complicated 2D and 3D meso-scale fracture processes in concrete, using the random field theories (Yang et al., 2009; Su et al., 2010), direct generation and packing of aggregates (Wang et al., 2015; Caballero et al., 2008), and XCT image-based models (Ren et al., 2015). Other approaches, such as dynamically inserting CIEs only when a certain criterion is reached (Yu et al., 2008; Ruiz et al., 2001; López et al., 2007), can also be used to improve the computational efficiency.

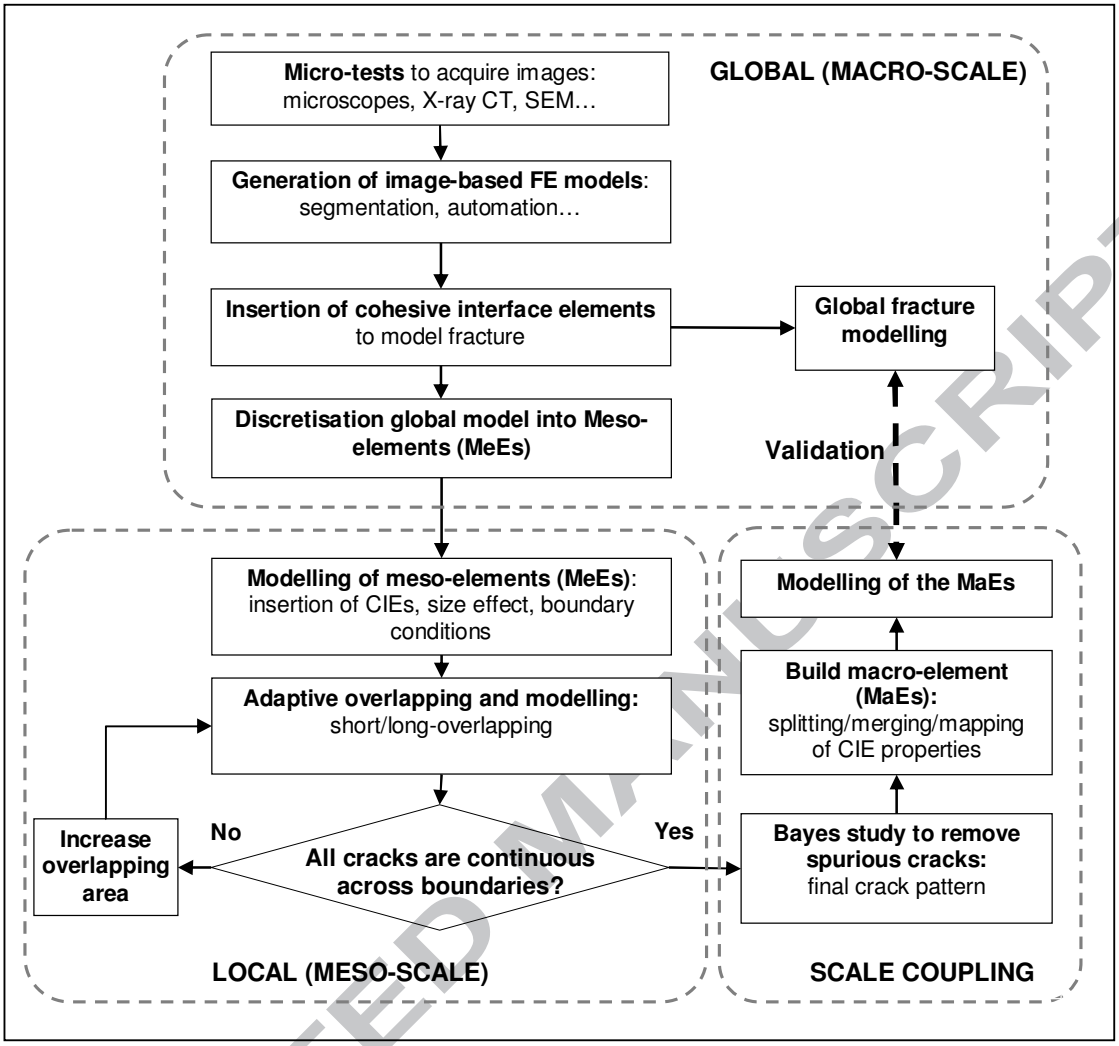


Figure 1: A two-scale coupling scheme for stochastic fracture mechanics.

The global domain is then divided into a grid with a number of non-overlapped rectangular meso-elements (MeEs), illustrated in Figure 3(a) as an example. For each MeE, its micro-structure is meshed with CIEs inserted and a nonlinear FE analyse is conducted under the same boundary condition. Size effect studies on the strength are carried out to determine a proper number and size of the MeEs. Two boundary conditions as illustrated on the first row in Figure 2, noted as B1 and B2 are necessary for the MeEs and scale transfer. The B1 is the uniaxial tensile condition and B2 is the shear condition necessary for the nonlinear integration of tractions for the global CIEs. At the global scale, two types of external uniaxial boundary conditions G1 and G2 are modelled, as illustrated in Figure 2(c) and 2(d), respectively. The popular G1 condition tends to result in the pure mode-I fracture mode, while the G2 allows core rotations and may lead to two main cracks. The G2 condition has also been increasingly used (Park and Paulino, 2012; Wang et al., 2015; Ren et al., 2015). In this paper, this was not implemented as a boundary condition for the scale transfer.

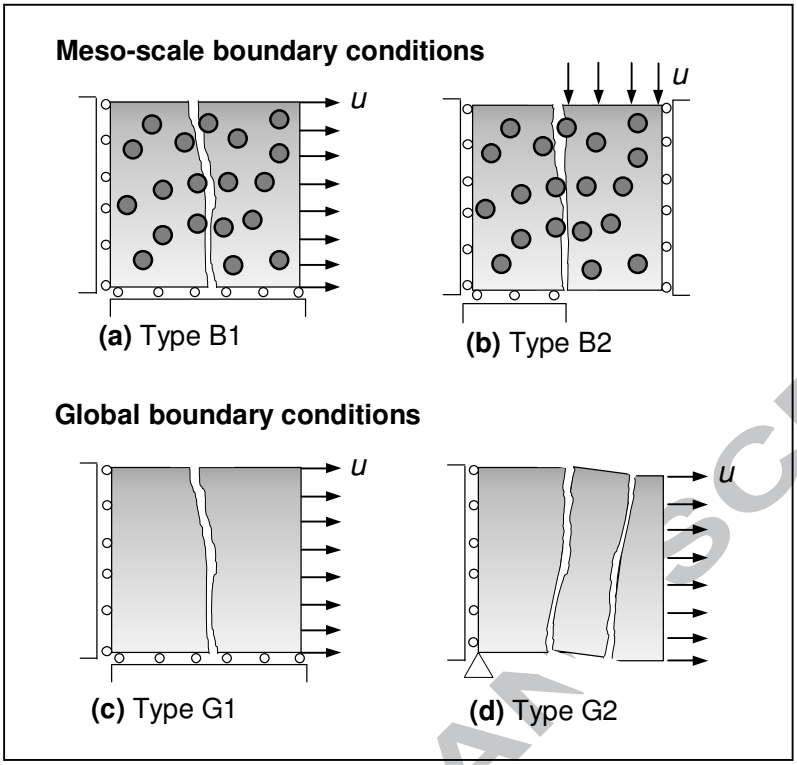


Figure 2: Boundary conditions for meso-scale modelling and for the global multi-scale models.

Figure 3(b) shows the crack paths from 16 independent nonlinear simulations of MeEs under the B1-X condition (X means horizontal and Y vertical direction).

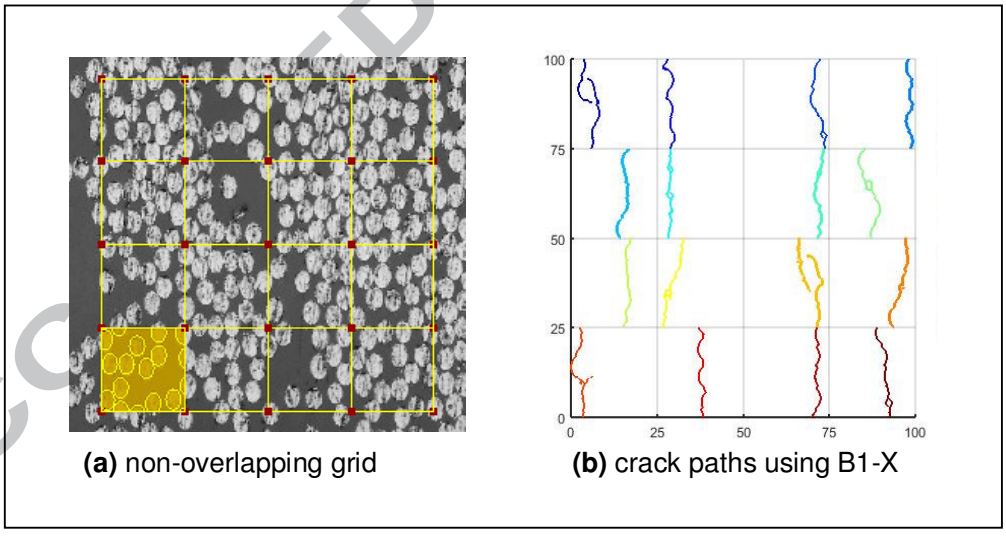


Figure 3: (a) A global domain discretised into 16 non-overlapped MeEs and (b) the modelled crack paths for each MeEs independently under B1-X boundary condition.

From Figure 3(b) it can be seen that not all the cracks across the MeEs' boundaries are continuous, indicating that deformation compatibility does not hold. This usually occurs when the non-overlapping grid is used. To improve the situation, an overlapping grid (Figure 4(a)) with the same

number but larger size MeEs is then designed. All the MeEs are again modelled independently. The resulting crack paths are shown in Figure 4(b). It can be seen that most of the crack paths now cross the MeEs' boundaries continuously, indicating improvement of the overlapping grid over the non-overlapping grid (Figure 3(b)). Subsequent overlapping grids with larger MeEs can be further designed and modelled if necessary. Figure 4(c) shows the final crack paths which nearly seamlessly cross all the MeEs' boundaries. This algorithm of using adaptively size-increasing overlapping grids works because the larger the overlapping regions become, the better deformation compatible support is provided.

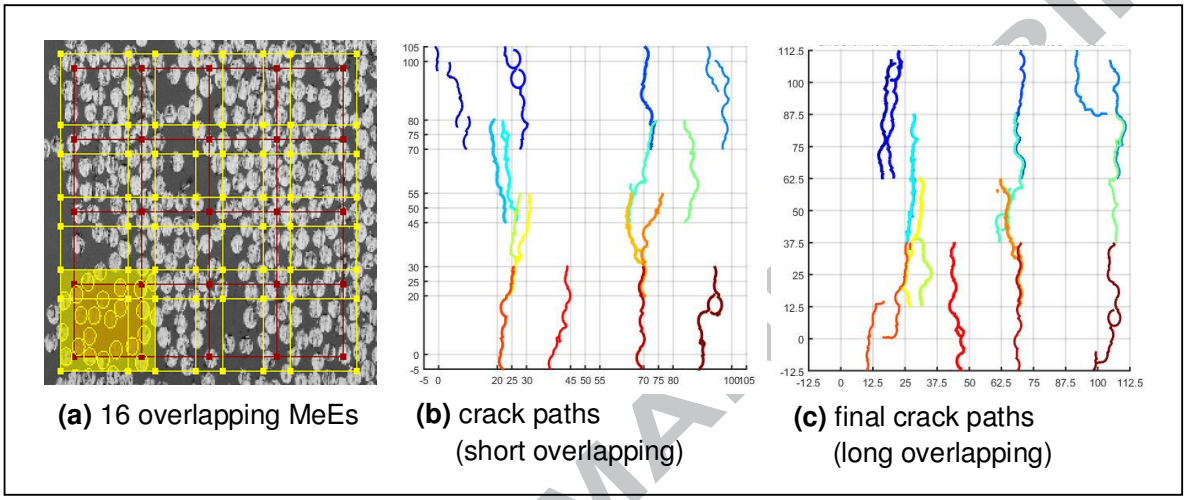


Figure 4: Overlapping MeEs and crack paths under B1-X boundary.

2.2. Bayesian inference model for overlapping MeEs with non-matching crack paths

Because the MeEs are separately modelled and the deformation compatibility cannot be rigorously ensured, some non-physically overlapped crack paths may be predicted, as seen in Fig. 4. To improve this situation, the Bayesian inference model is applied to filter the crack paths and preserve the most critical cracks only for subsequent scale transfer.

The well-known Bayesian inference method is

$$P(C|E)=\frac{P(E|C)}{\sum_m P(E|C_m)P(C_m)}\cdot P(C) \qquad \text{Equation 1}$$

Assume that the fracture process generates independent crack events E under similar boundary conditions but with unknown probability distribution. Equation 1 updates the degree of matching of a kernel crack path with m number of neighbouring cracks C_m (see Fig. 5). For each non-boundary kernel MeE, for example, there are 8 adjacent MeE each with a crack path so m=1~8. Assume that the denominator in Equation 1 can be represented by equally distributed probabilities $P(C_m)$ with known values for all adjacent half-overlapping and quarterly-overlapping MeEs with $\sum_{m=8} P(C_m)=1$. The probability $P(E|C_m)$ (m=1~8) remain fixed over the cluster C_m. It can be seen that Equation 1 updates the prior probability $P(E|C)$ to posterior probability $P(C|E)$ given a target probability $P(C)$. The $P(E|C)$ can be then used as a threshold to filter the cracks.

In the present paper the probabilities $P(E|C)$ and $P(C)$ are calculated as a surface/path integral, which can be reduced to a discrete number of key points at known locations as

$$P(C) = \int_C f(H_k) dH_k \qquad \text{Equation 2}$$

where the crack path is parameterised by a function $f(H_k)$ where H_k are key points on the crack path C. Alternatively Equation 2 can be calculated using the pixels forming the crack path in the images.

Fig. 5(a) illustrates two types of grid-free cracks: overlapping domain cracks and non-overlapping domain cracks when short overlapping is used. If the overlapping length is equal to or larger than half the MeE grid distance, the grid-free non-overlapping cracks merge into grid-free overlapping cracks. This becomes valid for all the non-boundary MeE due to full overlapping length. If the overlapping length exceeds half of the MeE grid distance, double overlapping regions appear. Due to the complexity in multiscale transfer, this study only deals with single overlapping grids of full-length, which bounds the long overlapping criteria to half the length of MeE size.

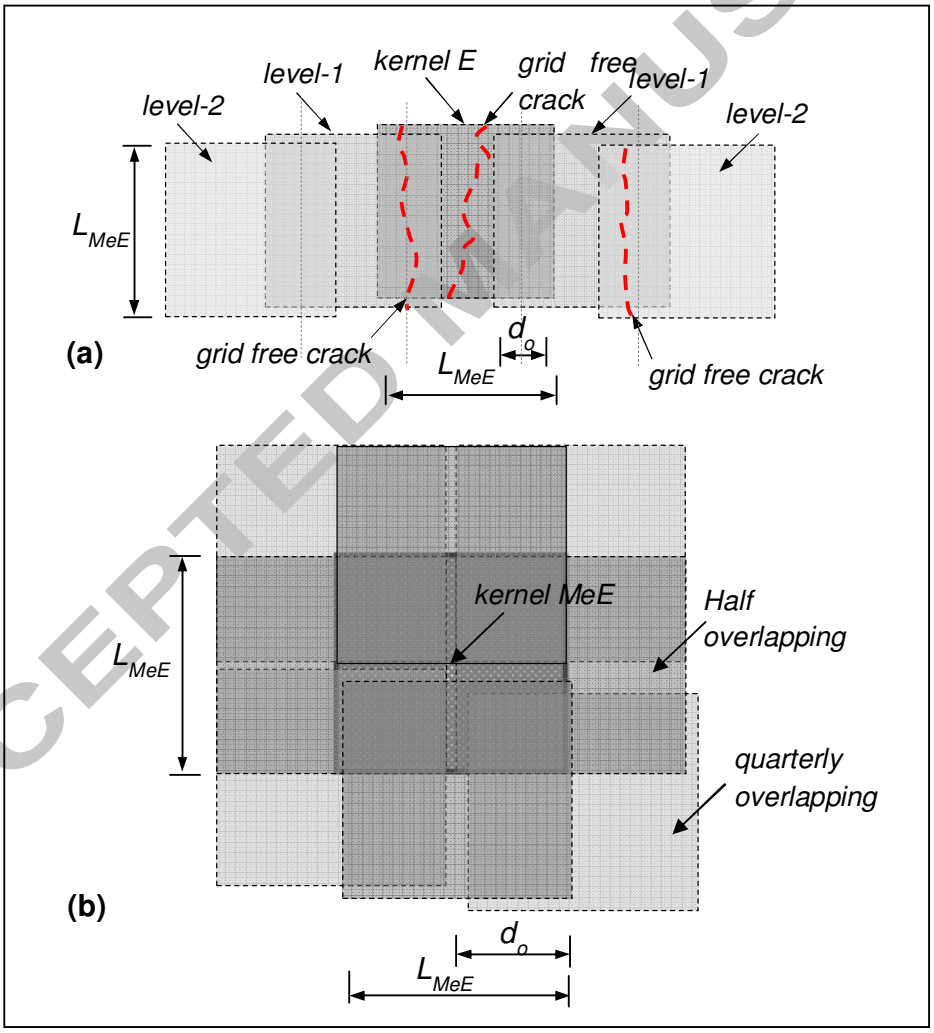


Figure 5: Illustration of crack sampling cases for the overlapping windows concept used with Bayesian criterion in Equation 1. The dark-shaded elements represent four half overlapping elements and soft-shaded elements represent four quarterly overlapping elements.

To quantify the matching degree with less computational effort, only two key points are initially used, i.e, the points of a crack path intersecting the non-overlapping grid. These points are suitable for any overlapping MeE sizes but may be unsuitable when multiple boundary conditions are simultaneously used. Therefore, the general case is implemented using matrices based on pixel metrics. Fig. 6 shows the implemented procedure and the results on an MeE using the half overlapping criteria.

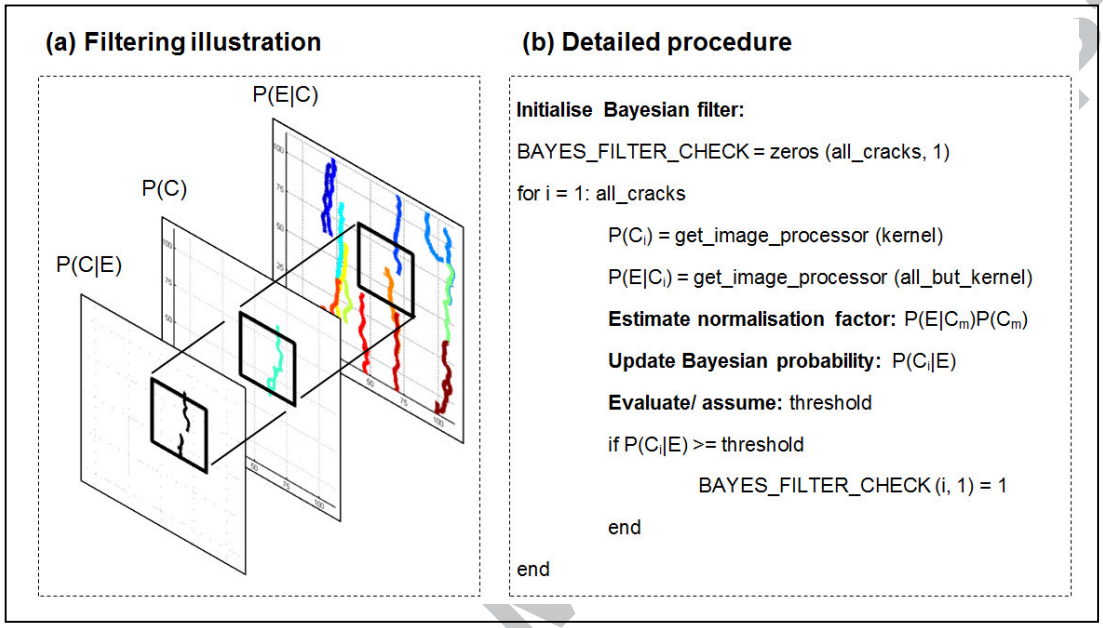


Figure 6: Bayesian filtering procedure and threshold probability on a typical non-boundary MeE

Fig. 7 shows the crack paths after using the Bayesian filtering method for two cases using the short and long overlapping grids.

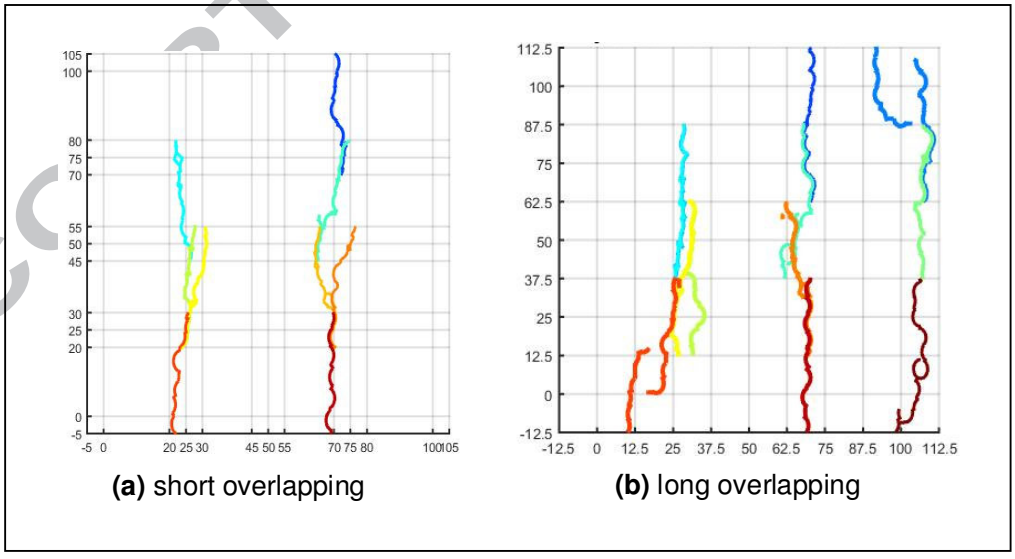


Figure 7: Filtered overlapping MeE crack path data after using the Bayesian inference model corresponding to Figure 4(b) & (c).

2.3. Construction of macro-scale elements (MaEs)

In general, short and long overlapping can be used as illustrated in Figure 8. This depends on the size establishment as for the MeE overlapping grid and are limited to single overlapping regions in this paper. This means that the size of an overlapping MeE cannot be larger than twice the MeE non-overlapping grid size (see Figure 3 (a) & Figure 4 (a)). In this paper, it is found that for defective materials, the short to long overlapping method may be used. For less defective materials, non-overlapping as for elastic studies or short overlapping discretisation in fracture mechanics may be used.

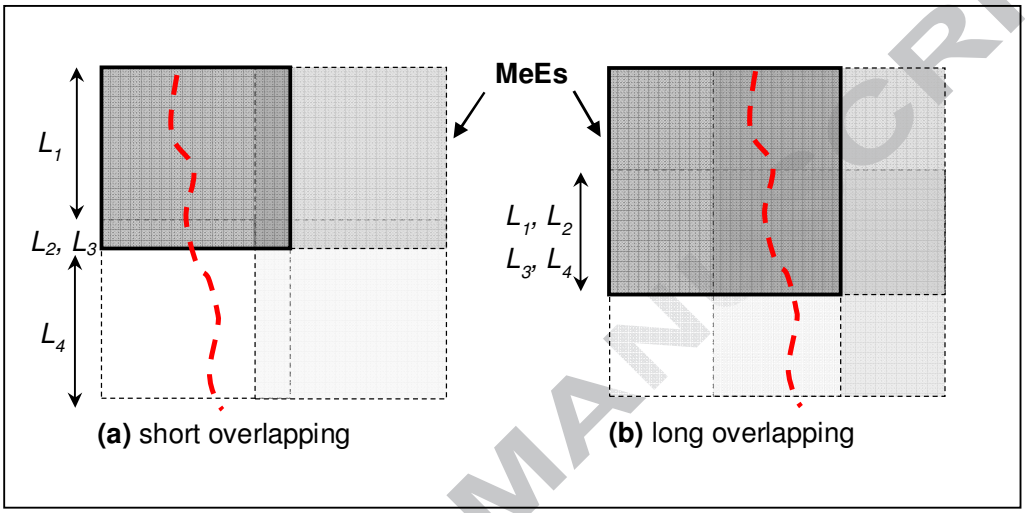


Figure 8: Illustration of (a) short to (b) long overlapping criteria. The colour intensities represent distinct MeE windows that are overlapped. The average fracture properties from corresponding crack path lengths denoted with L_i, where i is the corresponding overlapped crack path, are integrated piecewisely and transferred to the macro-scale by inserting cohesive elements.

After the optimised crack paths are found, macro-elements (MaEs) are constructed. The final meso-scale crack paths and non-overlapping grids are naturally used as guidelines to discretise the domain into a macro-mesh with a number of MaEs. In the macro-mesh, the crack paths are also modelled by CIEs whose softening constitutive laws are piecewisely mapped from the meso-CIEs, so that energy conservation is ensured in the scale transfer (see Figure 9).

The macro-scale mesh discretisation includes the intersection nodes, the assembly of the MaE model and the insertion of macro-cohesive interface elements. Figure 9 sketches two possible mesh models. The crack paths are identified based on the scalar degradation parameter d and the energy dissipation rate per unit volume of damage. In the case of short overlapping MeEs, the MaE nodes are at the intersections of the crack paths that cross cut the non-overlapping grids. The example illustrated in Figure 9 uses the B1-X boundary conditions. Two operations are used in the construction of the macro-meshes: merging and splitting (see Figure 9 (a) & (b)). These operations are generally used to reduce the integration order of the macro-model. However, they also become necessary when the crack path is very close or intersecting the non-overlapping grid. Their implications on the energy mapping rules are discussed in next section.

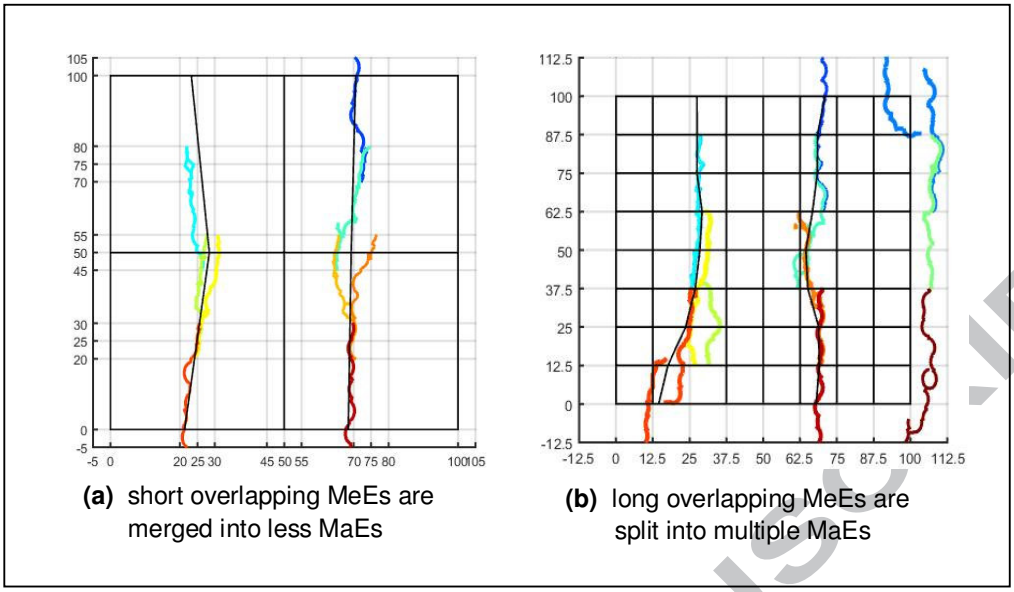


Figure 9: Illustration of two possible macro-scale meshes. The crack paths are used as boundary constraints during the MaE mesh generation. The cracks are also represented by cohesive interface elements (CIE) in the global model.

In fact, many other mesh models can be created at global length scale once the most critical crack paths available from the detailed simulations. For example, it is possible to build reduced integration meshes, either by using triangular or non-structured quads or by keeping details of the heterogeneous meso-scale mesh with coarser macro-elements and only inserting cohesive elements or enrichment nodes where necessary. However, there are some sources of uncertainty associated with the above strategy, for example:

- the centreline of the deformed cohesive crack MaEs may not be the same as the initial zero-thickness crack path which is used to build the MaE elements; this is due to the reversing of damage process for crack identification paths which were obtained using prescribed boundary conditions based on the stochastic element cluster assumption.
- the superposed straight edges of MaE elements in reduced order strategies may not be in line with the true integration crack paths which have a certain tortuosity; however, it is a reasonable assumption when using invariant crack paths to MeE boundary effects. This strategy can be used to reduce the global model size considerably.
- using more complicated mixed-mode boundary conditions (BCs) and highly non-homogeneous distributions of defects, two dominant crack failure modes can appear which may be tackled with a rotational centre. This imposes geometrical non-linearity in the macro-cohesive model which was not tested in this paper.

2.4. Energy mapping rules

After the macro-mesh is constructed, the traction-displacement softening curves of macro-CIEs need to be mapped from the meso-CIEs at the same position to ensure energy conservation.

The conversion of stress displacement curves in fracture energy is done according to the assumption in (Hillerborg et al., 1976) which defines the fracture energy required to open a unit area of crack by:

$$G_f = \int_{u_0}^{u_f} \sigma_{f,u} du \qquad \text{Equation 3}$$
$$\sigma_{f,u} = \sigma(u)$$

where σ_f and u_0 are the ultimate strength and displacement at the onset of fracture.

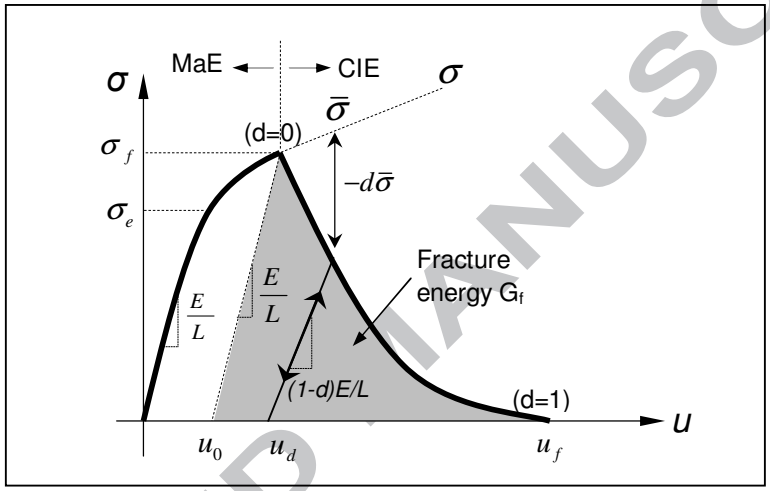


Figure 10: Illustration of the stress-displacement curve for the macro-scale bulk material properties. $\bar{\sigma}$ is the stress due to no damage, d is the scalar degradation variable and L is the characteristic length (Simulia/Abaqus).

For accurate integration across the scale transfer, L in Figure 8 can be represented by the actual crack length. This is feasible when appropriate scale discretisation and mesh order reduction strategy is considered. To prove the suitability of the new multi-scale concept, ABAQUS/Explicit is used in Section 3. In general, the estimated fracture energy from the collection of detailed MeE integration domains is consumed at macro-scale by the opening of macro cohesive elements. The method can be implemented by using both non-overlapping and overlapping adaptive strategy as discussed above. However, because the strategy uses fixed or enlarged computational areas but multiple parallel windows, this particular concept can be used to capture localized fracture behaviour with lower single simulation computational effort compared with large full-size detailed models. Two other advantages of the overlapping approach in this paper are: the capability of coupling weak to relatively strong interface models and its fully anisotropic formulation for analyses of complicated materials. This means that larger MeE windows can be adaptively employed for modelling materials with various degrees of flaws and fully anisotropic properties at any length scale. A case study for CFRP materials is presented in Section 3.

Different energy mapping rules are derived for merging and splitting operation respectively, considering both short and long overlapping.

2.4.1. Merging & splitting approach (MSA)

MSA is an edge oriented multi-scale cohesive crack model. The neighbour cluster of stochastic MeEs is either merged into larger size MaEs or split into several smaller size MaEs. The fracture energy dissipated is derived accordingly for the general cases based on average weighting factors of the overlapping regions as given in Equation 4 & Equation 5 for short and long overlapping respectively. Therefore, it should be noted that before using these two mapping rules, the existence of L_i factors contributing towards the macro-CIE should be first checked (see Figure 8).

$$G_{f,short} = \int_{u_0}^{u_f} \sigma_{f,L1} du + \frac{1}{2}\left(\int_{u_0}^{u_f} \sigma_{f,L2} du + \int_{u_0}^{u_f} \sigma_{f,L3} du\right) + \int_{u_0}^{u_f} \sigma_{f,L4} du \quad \text{Equation 4}$$

$$G_{f,long} = \frac{1}{4}\left(\int_{u_0}^{u_f} \sigma_{f,L1} du + \int_{u_0}^{u_f} \sigma_{f,L2} du + \int_{u_0}^{u_f} \sigma_{f,L3} du + \int_{u_0}^{u_f} \sigma_{f,L4} du\right) \quad \text{Equation 5}$$

The crack evolution at macro-scale can be also defined by means of a scalar degradation variable d which ranges between 0 and 1 (see explanation in Figure 10). The linear formulation of d can be expressed as:

$$d = \frac{u}{u_f} \quad \text{Equation 6}$$

Combining the Equation 4 & Equation 5 with Equation 6, the linear degradation variables become:

$$d_{short} = \frac{1}{3}\left(\frac{2G_{f,L_1}}{u_{f,L_1}\sigma_{f,L_1}} + \frac{G_{f,L_2}}{u_{f,L_2}\sigma_{f,L_2}} + \frac{G_{f,L_3}}{u_{f,L_3}\sigma_{f,L_3}} + \frac{2G_{f,L_4}}{u_{f,L_4}\sigma_{f,L_4}}\right) \quad \text{Equation 7}$$

$$d_{long} = \frac{1}{4}\left(\frac{2G_{f,L_1}}{u_{f,L_1}\sigma_{f,L_1}} + \frac{2G_{f,L_2}}{u_{f,L_2}\sigma_{f,L_2}} + \frac{2G_{f,L_3}}{u_{f,L_3}\sigma_{f,L_3}} + \frac{2G_{f,L_4}}{u_{f,L_4}\sigma_{f,L_4}}\right) \quad \text{Equation 8}$$

The degradation rates d can be also expressed exponentially. To ensure that the total energy dissipation on softening equals the total fracture energy G_f, the following expressions are derived:

$$d_{short} = 1/3 d_1 + 1/6(d_2 + d_3) + 1/3 d_4 \quad \text{Equation 9}$$

$$d_{long} = 1/4(d_1 + d_2 + d_3 + d_4) \quad \text{Equation 10}$$

where d_i are the individual contributions of MeEs to the degradation of the macro-cohesive element as illustrated in Figure 8 and given by:

$$d_i = 1 - \exp\left(-\int_{u_0}^{u_f} \sigma_{f,u,Li} du \Big/ G_{f,Li}\right) \quad \text{Equation 11}$$

This ensures that the fracture energy is gradually consumed during each displacement increment u_{Li}.

2.4.2. Crack path decomposition approach (CDA)

The CDA is an explicit meta-model, in the sense that cohesive elements are inserted only where necessary, while the homogenized continuum elements MaEs are integrated correspondingly so as to preserve the individual phases or composite elasticity. The material properties of cohesive elements for the overlapped mesh can be avoided if similar crack paths are repeating. Therefore, the general energy mapping rule involving non-overlapping single cracks or single-split cracks that overlap such as in Figure 9 (b) reduces to:

$$G_{f,L_i} = \int_{u_0}^{u_f} \sigma_{f,u,Li} du \qquad \text{Equation 12}$$

where the linear degradation variables are:

$$d_i = \frac{2G_{f,L_i}}{u_{f,L_i}\sigma_{f,u,L_i}} \qquad \text{Equation 13}$$

and the exponential degradation variables are:

$$d_i = 1 - \exp\left(-\int_{u_0}^{u_f} \sigma_{f,u,Li} du \Big/ G_{f,Li}\right) \qquad \text{Equation 14}$$

in which *i* corresponds to the crack site which is replaced by a 4-noded cohesive element. L refers to the actual crack length which is replaced by macro-scale edges. Alternatively, an exact stress strain evolution can be used giving tabular inputs (Simulia/Abaqus).

2.5. Computer implementation and procedure

The above methodology is implemented in a number of computer programs which are integrated in a batch file for automatic simulations. The meshing and the extraction of traction-displacement curves of CIEs are done by Python scripts. Pre-inserting CIEs, the construction of MaEs and energy mapping are implemented in MATLAB codes.

The MeE overlapping series were computed on parallel CPUs. The CPU time on a desktop PC i7 – 2600 @3.40GHz with 8 cores was about 5 to 6 hours per simulation, while when using 48 cores per simulation on the CSF facility at University of Manchester, the average time was 45 min. The large validation models MeE size 100μm were also simulated using 48 cores and lasted about 14 hours each.

3. CASE STUDY

This section presents the results of a case study applying the proposed multi-scale method to a three phase material (fibre, matrix and interface) made of carbon fibre reinforced polymer (CFRP).

3.1. Geometrical surveying of a CFRP ply

A geometrical survey was carried out first within the thickness of a ply to investigate the volume fraction variations (see Figure 11). The searching approach is similar to the one originally used for aluminum alloys (Graham and Yang, 2002) .

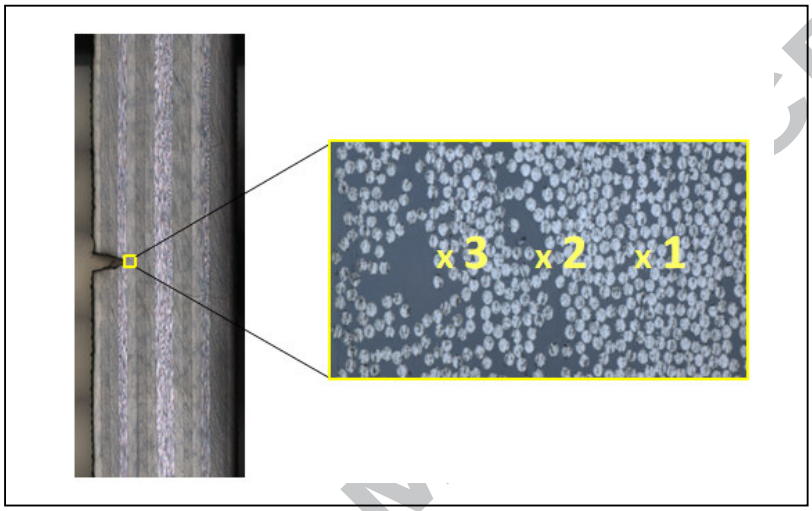


Figure 11: $+45^0$ ply in front of a notch tip (the numbers indicating positions of the centres of the imaging windows).

Figure 11 shows a micrograph extracted from a multi-layered CFRP beam with a V-notch. The images were acquired by image stitching from a laser confocal microscope. The micrograph covered the full thickness of a single layer of about 220µm. For this particular example, the fibres were elliptical shape of approximately 5µm of the shorter diameter because the ply ran at an inclined angle of +45 degrees. Three concentric square windows of different sizes were used to acquire the geometrical data. Figure 12 shows the typical acquisition windows at the mid position in Figure 11. Figure 13 summarises the surveying results.

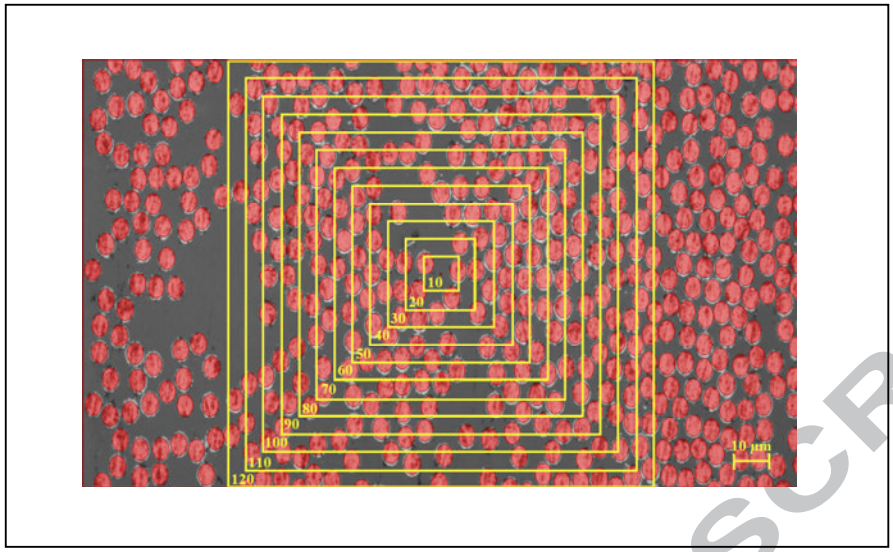


Figure 12: Surveying window sizes at position 2 in Figure 11.

Figure 13 also shows that within the ply thickness, the selection of the MeE window is morphologically important. The volume fraction statistics were found to be both position and size dependent. Apparently, the volume fractions converged when the window size was approximately 50µm^2, and the variation was more evident for small window sizes (10, 20 and 30µm^2) than larger sizes.

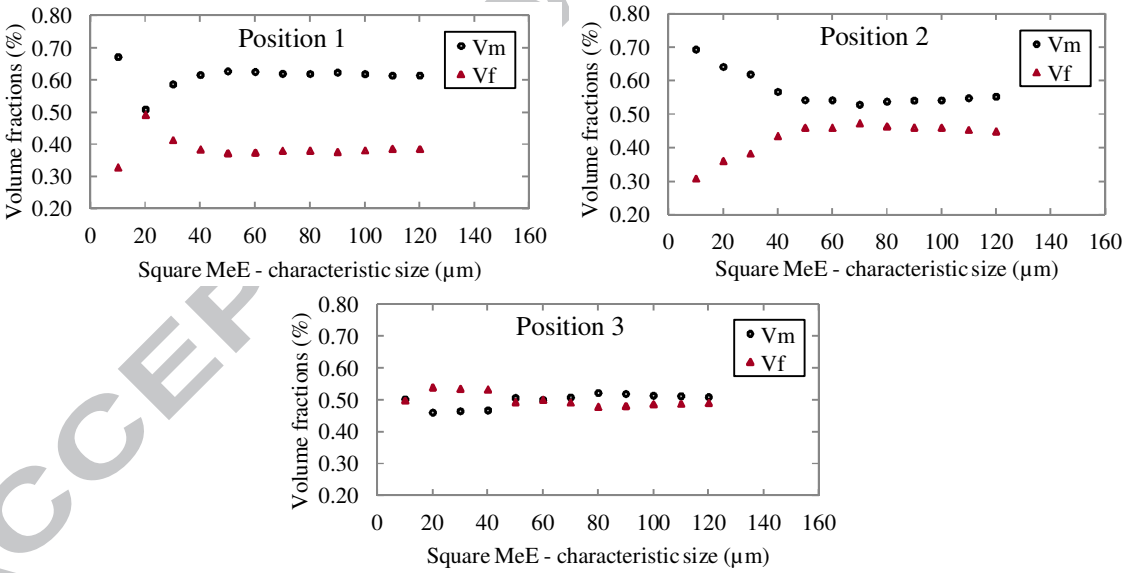


Figure 13: Variations of volume fractions of fibre (V_f) and matrix (V_m) with increasing MeE window size at different imaging positions for a carbon/epoxy ply +45°.

3.2. Material properties

Table 1 gives the main material parameters and the volume fractions used in the simulations. For simplification, it was assumed that the shear strength and fracture energy of each individual phase were equal to their normal ones respectively (Cid Alfaro et al., 2010; Vaughan and McCarthy, 2011; González and Llorca, 2007; Yang et al., 2009). Four types of fibre-matrix interfaces with different strength and fracture energy, namely, poor, weak, strong and perfect, were modelled here.

Table 1: Material properties and main modelling parameters.

	Carbon fibre	Epoxy resin	Interface
Elastic modulus E (GPa)	85	3.35	3.35
Poisson's ratio ν	0.22	0.35	0.35
Traction strength $t_n=t_s$ (MPa)	200	50	15 (poor), 25 (weak), 40 (strong), 50 (perfect)
Fracture energy $G_{fn}=G_{fs}$ (N/mm)	200E-03	50E-03	15E-03 (poor), 25E-03 (weak), 40E-03 (strong), 50E-03 (perfect)
Volume fraction (%)	30-60	70-40	-
Density (kg/m^3)		1500	

Table 2: Image based MeE simulations with modelling parameters.

Reference	Changing Parameters			Number of simulations per series
	Interface		Boundary conditions	
	$t_n=t_s$ (MPa)	$G_{fn}=G_{fs}$ (N/mm)		
MeE_10_100_W	25	25E-03	G1-X, G1-Y, G2-X, G2-Y	40
MeE_10_100_S	40	40E-03	G1-X, G1-Y, G2-X, G2-Y	40
MeE_16x25_P	15	15E-03	B1-X, B1-Y, B2-X, B2-Y	64
MeE_16x25_W	25	25E-03	B1-X, B1-Y, B2-X, B2-Y	64
MeE_16x25_S	40	40E-03	B1-X, B1-Y, B2-X, B2-Y	64
MeE_16x25_I	50	50E-03	B1-X, B1-Y, B2-X, B2-Y	64
MeE_16x35_W	25	25E-03	B1-X, B1-Y, B2-X, B2-Y	64
MeE_16x35_S	40	40E-03	B1-X, B1-Y, B2-X, B2-Y	64
MeE_16x50_W	25	25E-03	B1-X, B1-Y, B2-X, B2-Y	64
MeE_16x50_S	40	40E-03	B1-X, B1-Y, B2-X, B2-Y	64

Note: The first number of all series is the number of MeE windows. The second number in the MeE_10_100 series is the largest computed window in increments of size 10µm. The second number in the MeE_16x series is the overlapping grid size based on a non-overlapping grid base of size of 25µm.

3.3. Finite element modelling

The heterogeneous discretisation was performed using an in-house MATLAB code based on the pre-inserting cohesive element method in (Yang et al., 2009). At meso-scale, the finite element meshes consisted of triangular CPS3 solid elements and 4-noded cohesive interface elements COH2D4 in ABAQUS with linear traction-separation laws. The explicit dynamic solver with displacement control was used with adequate loading time for quasi-static loading condition.
Table 2 summaries the image-based simulations carried out with the boundary conditions and interface properties. All the other material properties used are shown in Table 1.

3.4. Validation of meso-scale modelling

Validation of the meso-scale modelling was conducted by comparison with results in (Cid Alfaro et al., 2010). An artificially created meso-scale window of S-glass fibre (30%) and epoxy resin was modelled. A good agreement was found in terms of stress contour plots and crack propagation patterns. Figure 14 shows the comparable fracture behaviour. The cracks are represented by the cohesive interface elements with damage index $D \geq 0.9$.

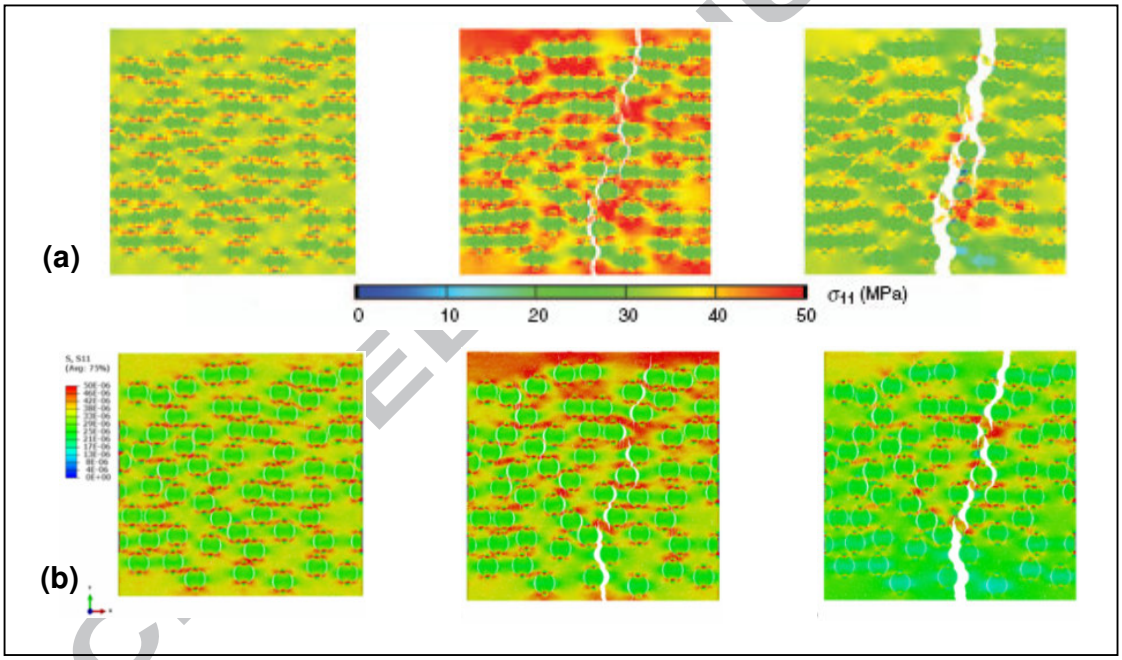


Figure 14: Comparison of stress contour plots on window size 125 µm and fibre volume fraction 30% at three different load steps: (a) results in (Cid Alfaro et al., 2010), (b) the present study.

To further demonstrate the importance of the interface material strength, a MeE of size 50µm^2 was simulated using the four interfacial types in Table 1. Figure 15 shows that the interface properties can greatly influence the meso- cracking mechanism which has important crack propagation effects on the macro-scale fracture. This is because local stress concentrations may potentially divert crack propagation for certain global boundary conditions. Experiments are thus necessary to examine the interface properties between fibres and the matrix and their mechanical effects.

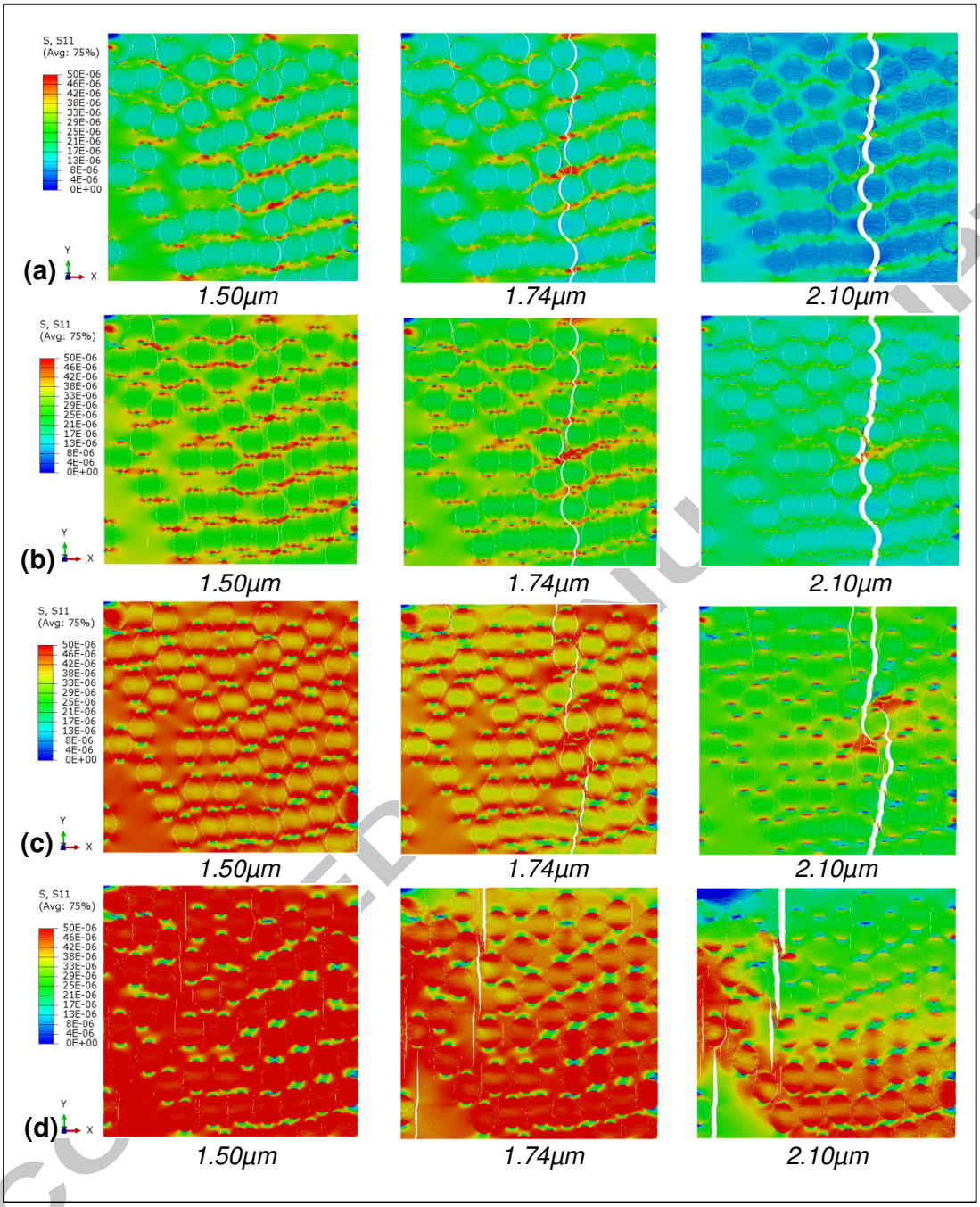


Figure 15: Stress contours at three load steps from MeE number 3 of size 50μm^2 using (a) poor, (b) weak, (c) strong and (d) perfect interface properties.

Boundary effects also influence crack initiation and propagation and therefore the overall dissipated energy. Figure 16 shows a two-crack dominant failure in the MeE size 100μm^2 image-based model using the G2-X type boundary conditions from Figure 2. In general, such effects were not observed on the other boundary types; thus only B1 type that is identical to G1 boundaries was further used in the scale coupling.

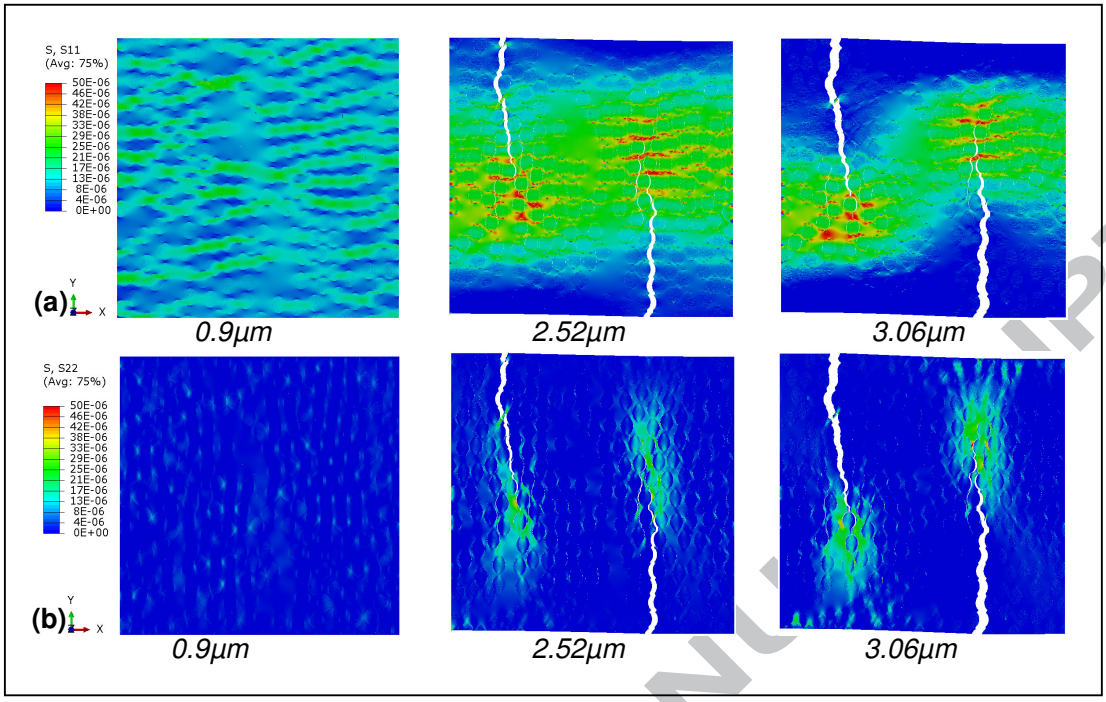


Figure 16: Stress contours at three loading steps using G2-X boundaries and weak interfacial properties showing a two dominant cracks failure in MeE size 100μm^2.

3.5. MeE size effects

Simulations for different MeE sizes, in increments of $10\mu m^2$ up to $100\mu m^2$, with centre at position 2 in Figure 11, were carried out to investigate size effects on crack paths and load-displacement curves. To avoid problems of localisation and mesh interface disclosure for zoom-out resolution, surface partitions were conducted so that the same mesh in smaller MeE is contained in larger ones (see Figure 17).

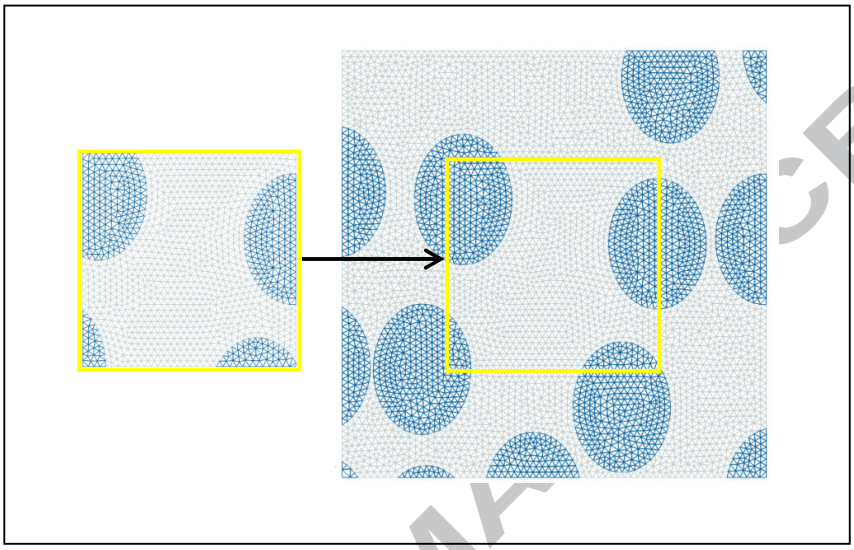


Figure 17: Image-based MeEs of size $10\mu m^2$ and $20\mu m^2$ illustrating the concept of partitioned concentric windows by preserving mesh topology.

The boundary conditions are given in Table 2. It is interesting to see that from one size to another, dissimilar crack paths appear when using G1-X loading conditions (see Figure 18 and Figure 19). This is expected since image based models have random inclusions and defects. Also the window sizes are often not sufficiently large to achieve a more uniform failure mode. A typical set of stress-displacement curves and the equivalent stress-strain curves are shown in Figure 20(a) & (b) respectively.

Figure 20(b) shows that a more sudden failure occurs as the MeE size increases. A strong size effect of the peak stress was found especially when weak interface properties were simulated. For all types of interfaces, the size effect of the strength decreases when the window size is larger than 20 - 30µm which was set as a minimum size of the MeE required for the scale transfer (see Figure 21). This is consistent with the statistical analysis of volume fractions in Figure 13.

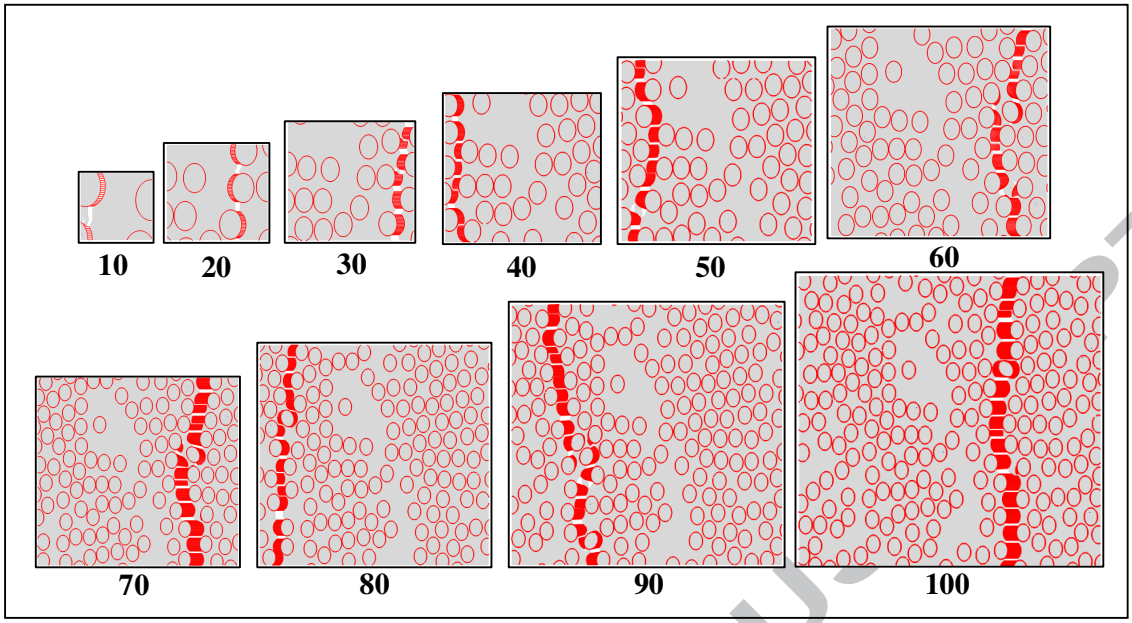


Figure 18: G1-X crack paths in different MeE sizes in μm for weak interface properties (the interface cohesive layer illustrated in red).

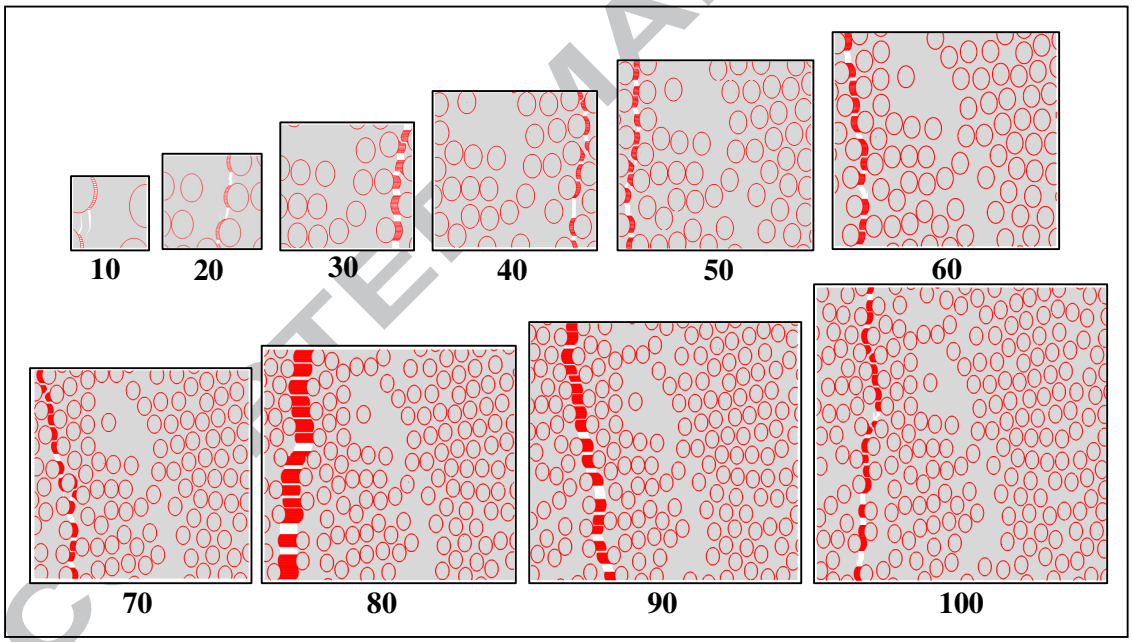


Figure 19: G1-X crack paths in different MeE sizes μm for strong interface properties (the interface cohesive layer illustrated in red).

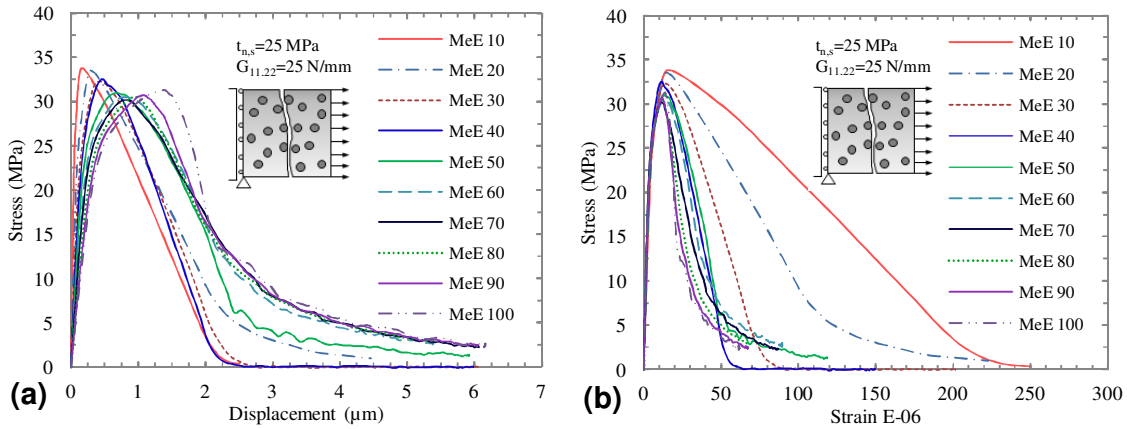


Figure 20: Stress-displacement and stress-strain curves for G2-X boundary conditions on MeE step $10\mu m^2$ to $100\mu m^2$ by using weak interface properties.

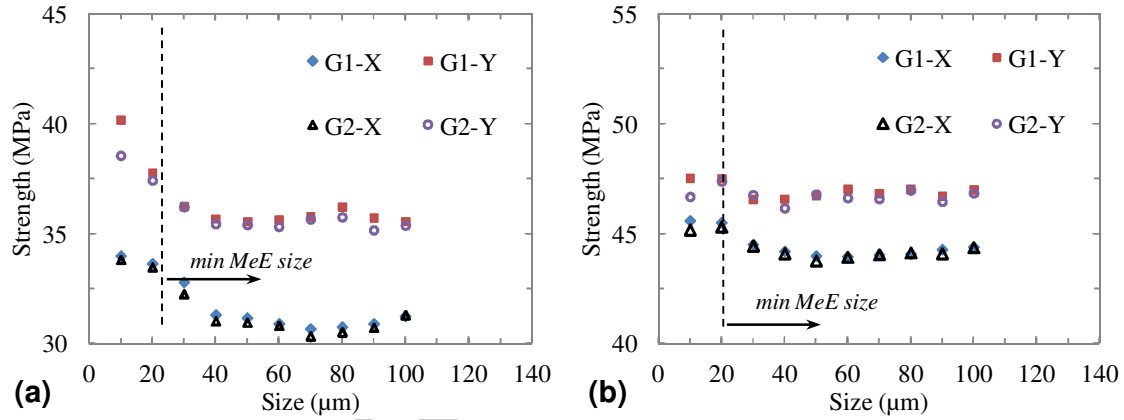


Figure 21: Size-strength results for MeE step $10\mu m^2$ to $100\mu m^2$ using (a) weak and (b) strong interface properties.

3.6. Stochastic MeE simulations

To investigate the effects of crack path bias and local orientation diversions in scale transfer, non-overlapping meso scale windows such as shown in Figure 3(a) were used in simulations first. The square MeE windows were simulated using the full set of boundary conditions in Table 2. The adaptive short to long-overlapping MeE grid discretisation was then used to tackle the problem of neighbouring crack paths bias. The idea is to share an overlapped support as to solve deformation compatibility. The concept may be also understood as an inverted overlapping limit. The upper bound corresponds to 2x computation of the same core MeE windows. This means that for long overlapping the integration must be limited by the distance between two neighbouring mass centres to avoid complicated double overlapping areas. This is useful particularly when strength and fracture energy variances are relatively small which holds in most composites with random distribution of features. Obviously, the short or narrow overlapping criteria will bring further computation savings as the computational areas are closer to the non-overlapping model. However, for both cases the main advantage is that the simulation of global large models can be parallelized to achieve high computational efficiency.

In Figure 22, the results for two main scenarios are presented as follows: (a,b,c) show weak interface crack paths and (d,e,f) show strong interface crack paths which were obtained using B1 boundaries on the orthotropic directions X and Y. Thus, given the random distribution of fibres, the interface layer between the fibres and matrix can be considered an intrinsic defect entity in this study. It can be seen that, when the non-overlapping method was used (a,d), the fracture paths rarely matched between the neighbouring MeE elements. When overlapping MeE windows were used, the matching rates were better in the case of weak interface properties (see Figure 22). In addition, there was a clear enhancement in crack site prediction compared with the non-overlapping results for cracks near edges (see and compare Figure 22 (a,b,c) & (d,e,f)). Figure 23 shows the variation of the predicted ultimate strengths with varying overlapping areas for different MeEs. Smaller strength variance with larger prediction errors were obtained for the strong interface models (see Figure 23 (b) MeE 12 & MeE 14).

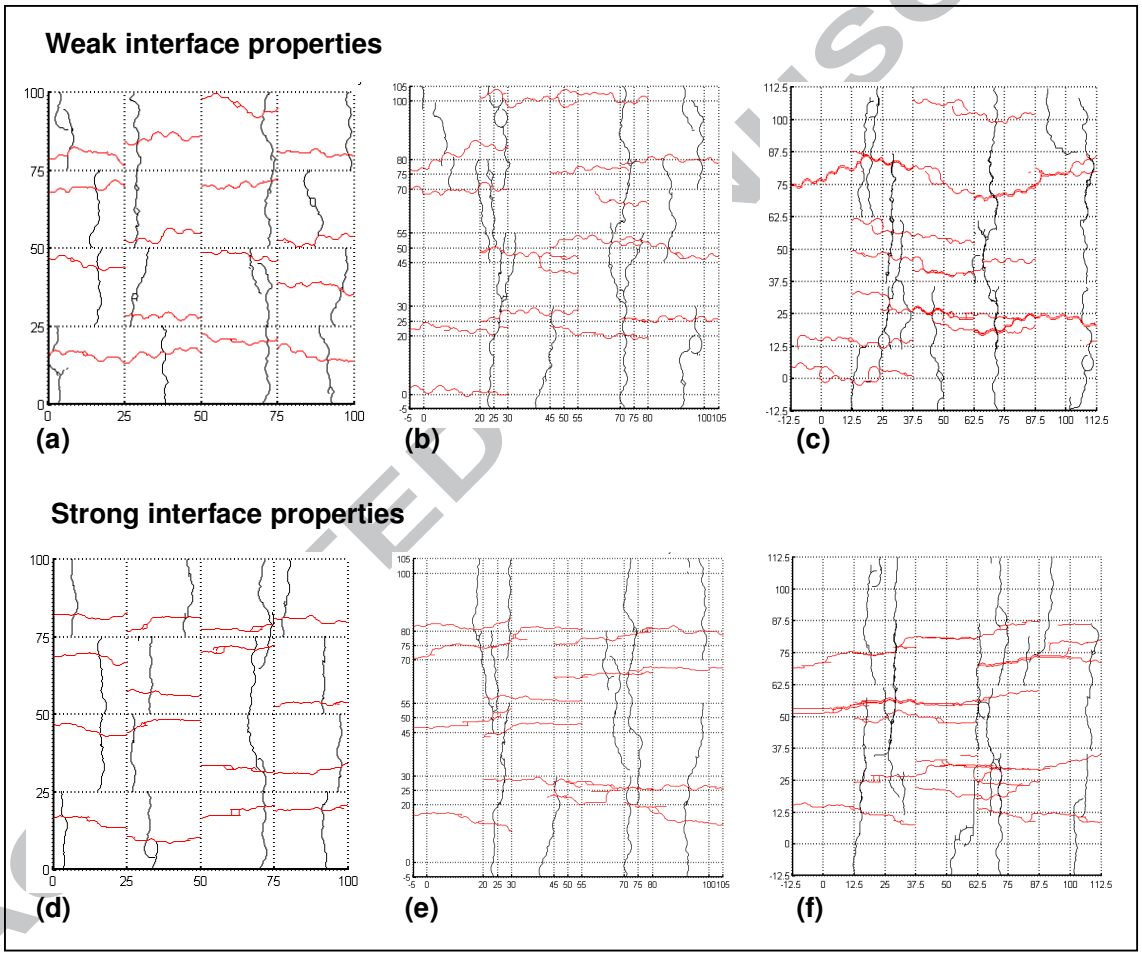


Figure 22: Crack paths for the adaptive size increasing non-overlapping MeE 16x25 (a,d) to short overlapping MeE16x35 (b,e) and long overlapping MeE16x50 (e,f). The two sets are from using weak interface properties (a,b,c) and strong interface properties (d,e,f).

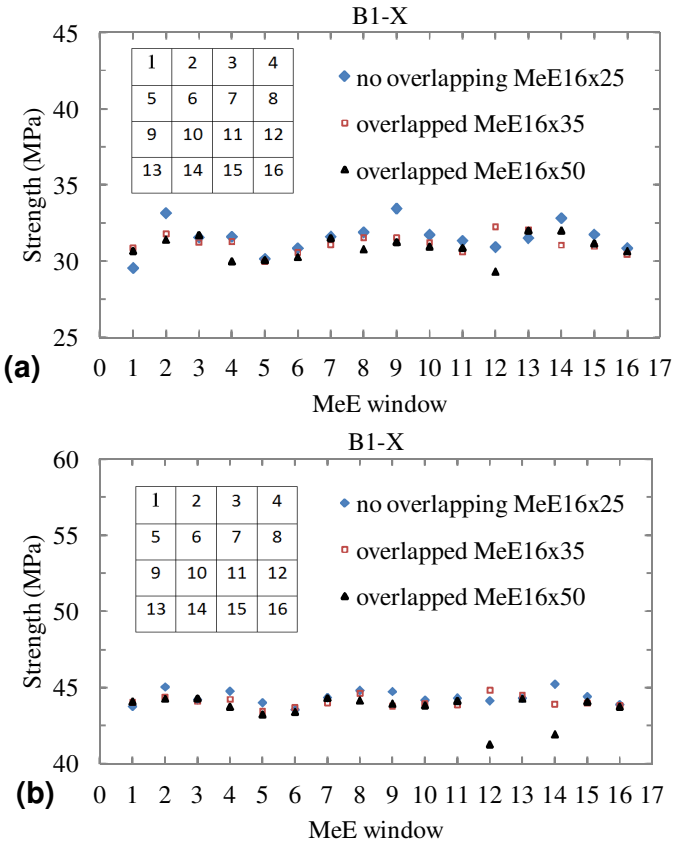


Figure 23: Variation of MeE strengths with different overlapping areas: (a) weak and (b) strong interface properties.

It should also be mentioned that for the scale transfer the B2-X and B2-Y boundary conditions in Figure 2 are necessary. The confidence domain for the softening part in such a case was limited to small sliding. This domain is equal to about 0.2µm which is approximately the size of one finite element used in the MeE simulations. A mesh example is shown in Figure 17.

In the following, the validation of the proposed multi-scale modelling method is established by comparing the detailed crack propagation results and energy dissipation mechanisms on both scales.

3.7. Comparison of meso and macro-crack propagation

This section elucidates the most convenient averaging approach to solving the scale transfer problem. The effective stiffness for each individual MeE was evaluated based on the results in Section 3.6. ABAQUS offers an orthotropic elasticity model which can be defined by a stiffness matrix. This matrix incorporates the effects of individual fracture modes on X and Y directions and the analytically combined effect of both (Simulia/Abaqus). A local coordinates system was also defined for the orthonormal directions X and Y.

Figure 24 & Figure 25 compare all the crack paths of uncoupled overlapping MeEs using B1 boundaries (shown in coloured lines) against the corresponding fully detailed MeE100 simulations (shown in black lines) using G1 and G2 cases on X and Y respectively.

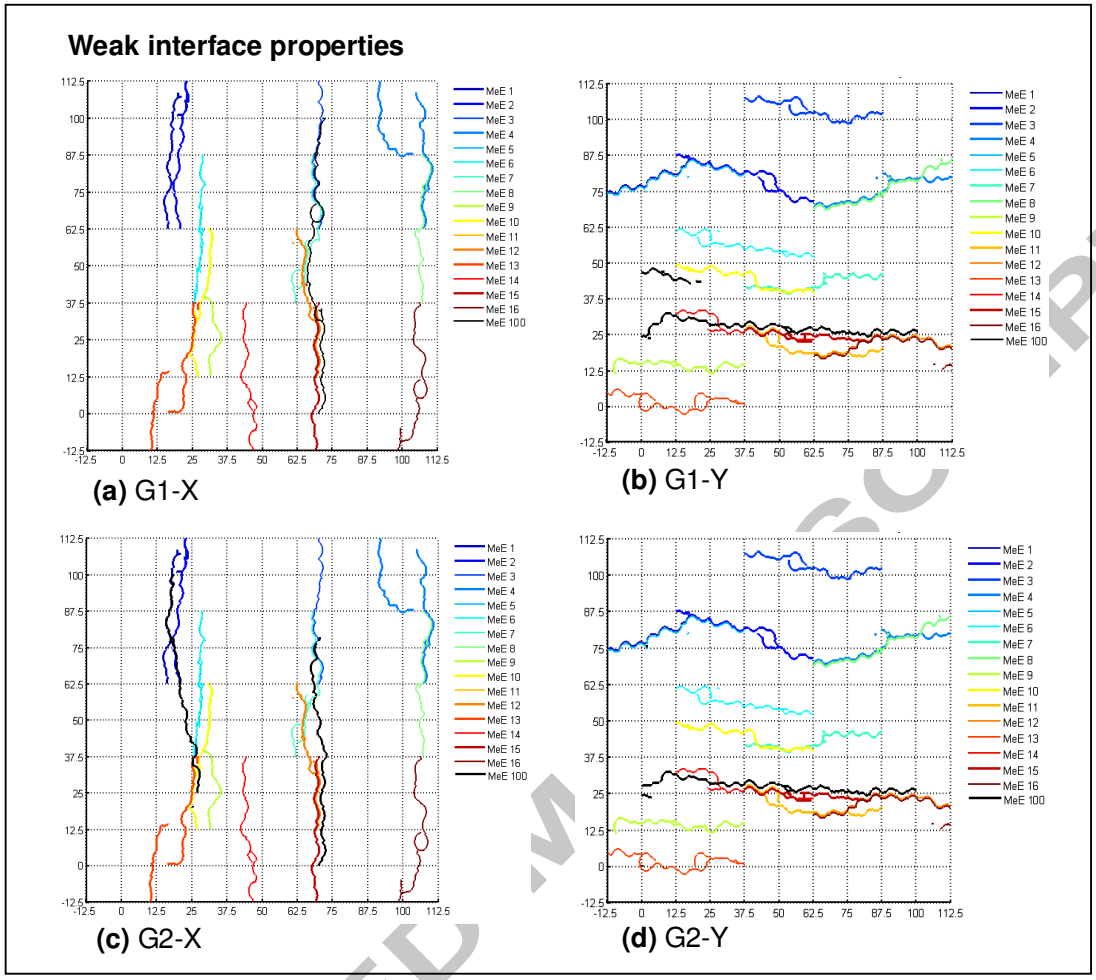


Figure 24: Comparison of separate crack paths between overlapping series MeE16x50 (shown in coloured lines) using B1 boundaries against full size MeE100 (shown in black lines): (a) G1-X; (b) G1-Y; (c) G2-X and (d) G2-Y for weak interface properties.

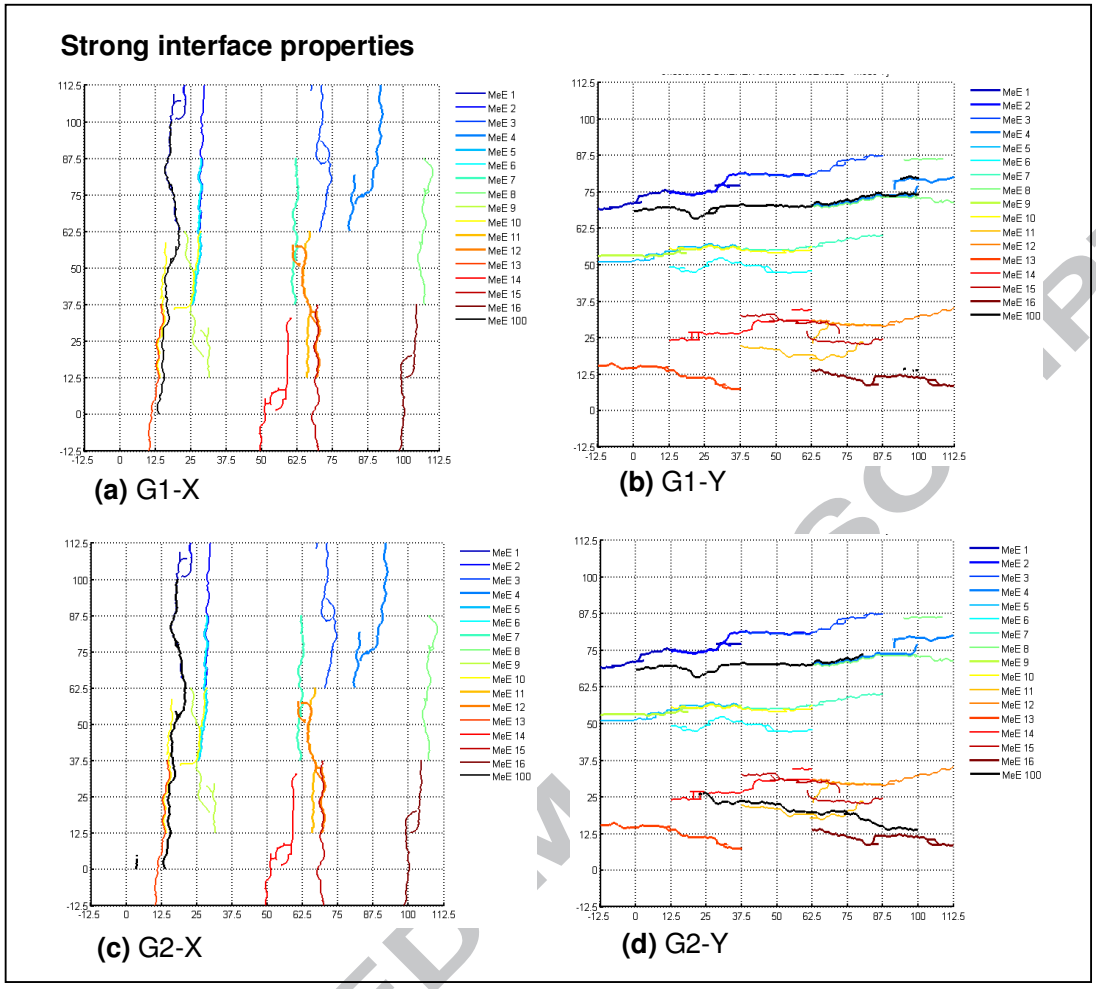


Figure 25: Comparison of separate crack paths between overlapping series MeE16x50 (shown in coloured lines) using B1 boundaries against full size MeE100 (shown in black lines): (a) G1-X; (b) G1-Y; (c) G2-X and (d) G2-Y for strong interface properties.

Figure 26 & Figure 27 compare the stress-displacement relationships of multi-scale models using overlapping series MeE16x50 and detailed MeE100 for weak and strong interface properties respectively. These models use reduced meshes and integrate material properties according to the methodology in Sections 2.3 & 2.4. The corresponding crack paths of both meso- and macro-simulations are also shown. It can be seen that the agreement was good when a single dominant crack path was modelled. If there are two distant crack paths, the matching locations could be established only in the case of weak interface properties (see the G2-X results in Figure 26 & Figure 27).

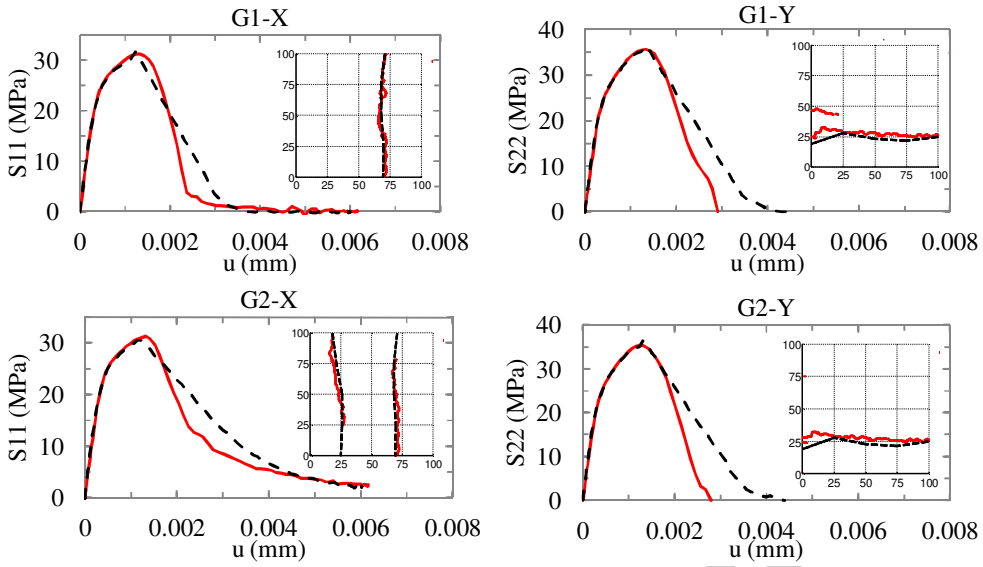


Figure 26: Nonlinear multi-scale results for weak interface properties using the overlapping concept vs fully detailed simulations (solid curve: detailed MeE100, dashed curve: using MeE 16x50 assembly).

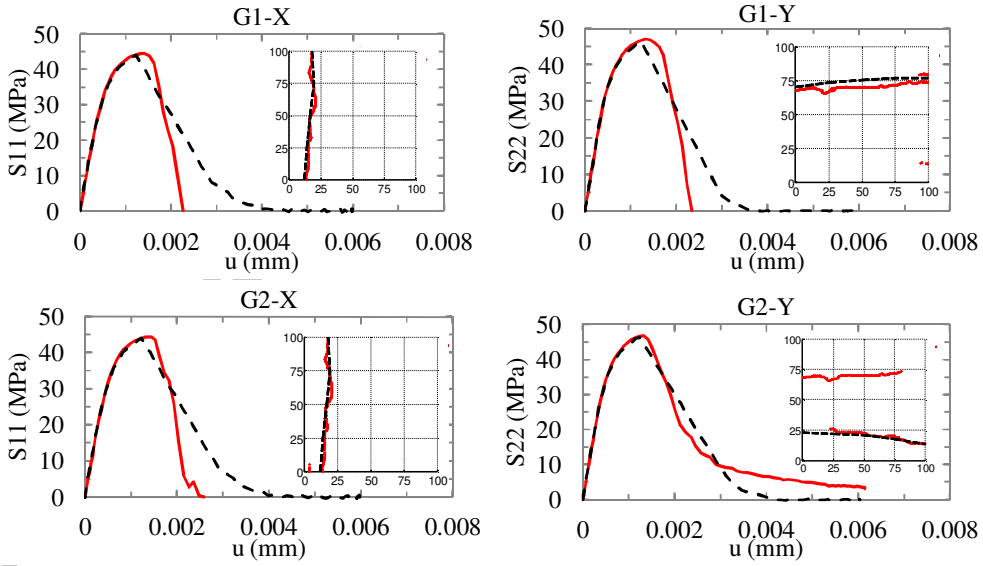


Figure 27: Nonlinear multi-scale results for strong interface properties using the overlapping concept vs fully detailed simulations (solid curve: detailed MeE100; dashed curve: using MeE16x50 assembly).

3.8. Application of the Bayesian inference

The Bayesian inference model in section 2.2 is then used to filter the non-matching crack paths in Fig. 24 and Fig. 25. The results are presented in Figure 28 & Figure 29, which show that highly nonlinear situations at global scale could be captured effectively by inserting macro-CIEs only where

necessary. The results are much improved with a good agreement especially for the two dominant crack paths situation in Figure 28 and Figure 29 using G2-X and G2-Y boundaries respectively.

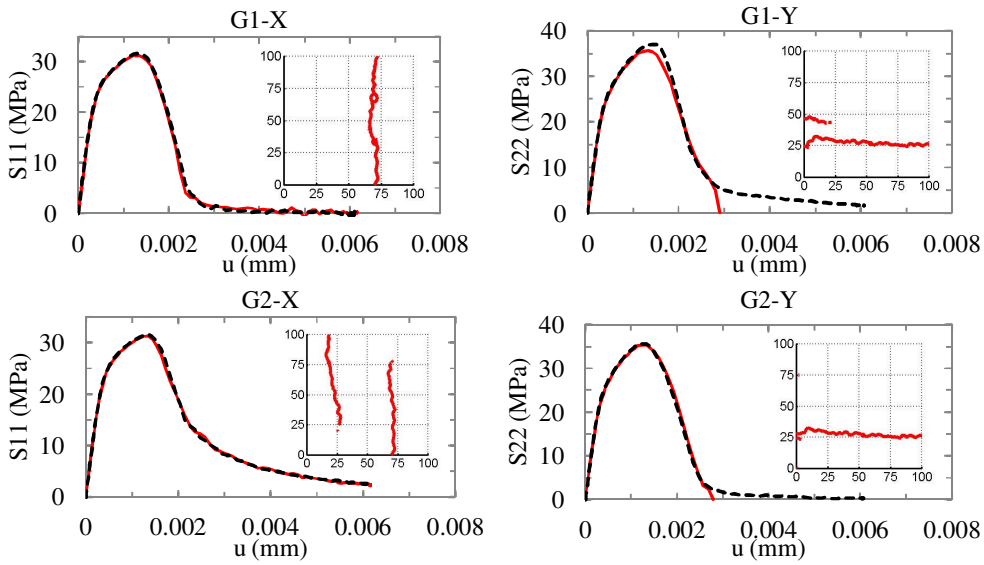


Figure 28: Comparison of simulation results using the multi-scale stochastic coupling strategy for weak interface properties. The solid curves represent the detailed geometry models (i.e. MeE size 100μm^2) and dashed curves are the Bayesian multi-scale models.

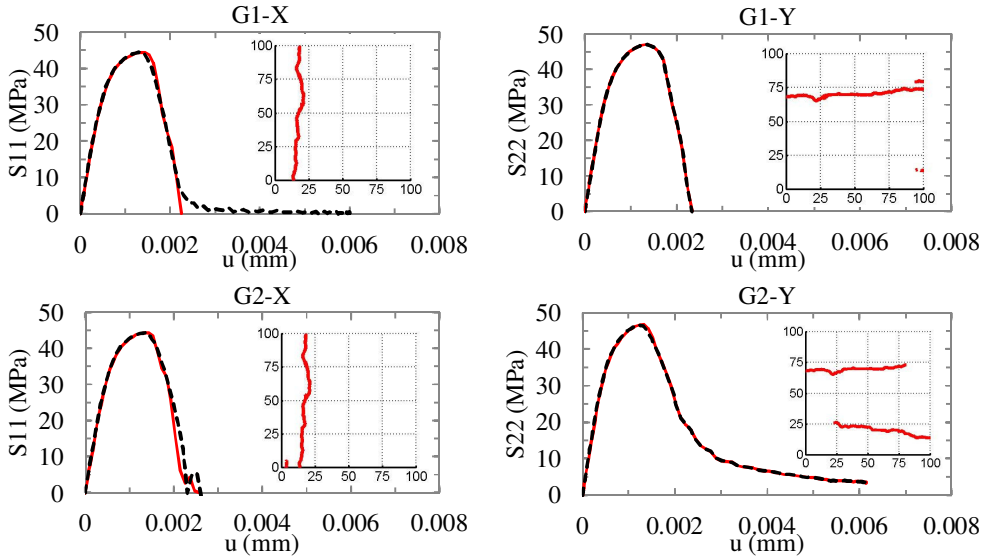


Figure 29: Comparison of simulation results using the multi-scale stochastic coupling strategy for strong interface properties. The solid curves represent the detailed geometry models (i.e. MeE size 100μm^2) and the dashed curves are the Bayesian multi-scale models.

4. CONCLUSIONS

This study has presented the development and validation, through a numerical case study, of a new multi-scale stochastic fracture mechanics framework (MsSFrM) for modelling cohesive fracture in heterogeneous materials. The following conclusions may be drawn:

(1) The selection of the meso-scale window sizes is highly related to the distribution of volume fractions. For accurate multi-scale couplings, the morphological continuity should be ensured by accurate mapping of the individual MeEs.

(2) Increasing the MeE size does not necessarily stabilise the fracture behaviour, thus invalidating the classical RVE approach. However, by choosing the appropriate MeE window size and resolving the boundary deformation compatibility via the overlapped windows concept, the crack bias effects can be minimized and therefore the scale transfer could be captured effectively.

(3) Energy mapping rules have been proposed for solving the scale coupling problem for different interface properties when the crack matching criteria is met. For defective materials, the short to long overlapping method may be used. For less defective materials, non-overlapping as for elastic studies or short overlapping can be used.

(4) The Bayesian inference model can be used to reduce the uncertainty of bias crack paths in the multi-scale transfer owing to its ability to update the state of belief using element cluster searches.

(5) The new multi-scale stochastic modelling method offers a more accurate modelling framework for multiphasic heterogeneous materials with various levels of defects. Heterogeneity features such as material defects, inclusions and voids can be incorporated at both scales.

(6) It should be noted that in the present method, all the MeEs are modelled independently under the same set of tensile and shear boundary conditions. This is not strictly valid even for the case studies under uniaxial tension in this paper. The results in terms of both final crack paths and load-displacement curves, however, are very good, probably due to the use of adaptively-increasing overlapping grids which may have ensured the deformation compatibility between adjacent MeEs to some extent. Research is much needed to integrate accurate algorithms for solving compatible boundary conditions of individual MeEs, such as those discussed by Oden et al (1999), so that the developed method can be applied to general situations.

Acknowledgements

This study is funded by a US Air Force EOARD grant (No. FA8655-12-1-2100) and an EPSRC grant (No. EP/J019763/1). We would like to thank Prof Costas Soutis at the Northwest Composite Centre in Manchester for providing CFRP materials, Dr William Bodel at Nuclear Graduate Research Group (NGRG) and C-Net for helping with acquisition of images and in-situ microscopy.

References

Aarnes, J. E., Krogstad, S. & Lie, K.-A. (2006) A Hierarchical Multiscale Method for Two-Phase Flow Based upon Mixed Finite Elements and Nonuniform Coarse Grids. *Multiscale Modeling & Simulation,* 5(2)**,** 337-363.

Belytschko, T. (2007) Research directions in computational and composite mechanics. *A Report of the United States National Committee on Theoretical and Applied Mechanics*.

Belytschko, T., Loehnert, S. & Song, J.-H. (2008) Multiscale aggregating discontinuities: A method for circumventing loss of material stability. *International Journal for Numerical Methods in Engineering,* 73(6)**,** 869-894.

Bosco, E., Kouznetsova, V. G. & Geers, M. G. D. (2015) Multi-scale computational homogenization-localization for propagating discontinuities using X-FEM. *International Journal for Numerical Methods in Engineering,* 102(3-4)**,** 496-527.

Caballero, A., Willam, K. J. & Carol, I. (2008) Consistent tangent formulation for 3D interface modeling of cracking/fracture in quasi-brittle materials. *Computer Methods in Applied Mechanics and Engineering,* 197(33-40)**,** 2804-2822.

Calo, V. M., Efendiev, Y., Galvis, J. & Ghommem, M. (2014) Multiscale empirical interpolation for solving nonlinear PDEs. *Journal of Computational Physics,* 278**,** 204-220.

Canal, L. P., González, C., Segurado, J. & Llorca, J. (2012) Intraply fracture of fiber-reinforced composites: Microscopic mechanisms and modeling. *Composites Science and Technology,* 72(11)**,** 1223-1232.

Carpinteri, A. & Chiaia, B. (1997) Multifractal Scaling Laws in the Breaking Disordered Materials.

Carpinteri, A., Chiaia, B. & Cornetti, P. (2002) A scale-invariant cohesive crack model for quasi-brittle materials. *Engineering Fracture Mechanics,* 69(2)**,** 207-217.

Chen, Y., Durlofsky, L. J., Gerritsen, M. & Wen, X. H. (2003) A coupled local–global upscaling approach for simulating flow in highly heterogeneous formations. *Advances in Water Resources,* 26(10)**,** 1041-1060.

Chung, E. T., Efendiev, Y. & Fu, S. (2014) Generalized Multiscale Finite Element Method for Elasticity Equations. *ArXiv e-prints,* (1408.5929).

Cid Alfaro, M. V., Suiker, A. S. J. & de Borst, R. (2010) Transverse Failure Behavior of Fibre-epoxy Systems. *Journal of Composite Materials*.

Coenen, E. W. C., Kouznetsova, V. G., Bosco, E. & Geers, M. G. D. (2012) A multi-scale approach to bridge microscale damage and macroscale failure: a nested computational homogenization-localization framework. *International Journal of Fracture,* 178(1-2)**,** 157-178.

Cusatis, G. & Cedolin, L. (2007) Two-scale study of concrete fracturing behavior. *Engineering Fracture Mechanics,* 74(1-2)**,** 3-17.

de Borst, R. (2008) Challenges in computational materials science: Multiple scales, multi-physics and evolving discontinuities. *Computational Materials Science,* 43(1)**,** 1-15.

Desmorat, R. & Lemaitre, J. (2001) A Two-Scale Model for Quasi-Brittle and Fatigue Damage. *In:* Jean, L. (ed.) *Handbook of Materials Behavior Models.* Burlington: Academic Press.

Dirrenberger, J., Forest, S. & Jeulin, D. (2014) Towards gigantic RVE sizes for 3D stochastic fibrous networks. *International Journal of Solids and Structures,* 51(2)**,** 359-376.

Efendiev, Y., Galvis, J., Li, G. & Presho, M. (2014) Generalized multiscale finite element methods: Oversampling strategies. 12(6)**,** 465-484.

Efendiev, Y., Hou, T. Y. & Ginting, V. (2004) Multiscale Finite Element Methods for Nonlinear Problems and Their Applications. 553-589.

Fish, J. (2011) Multiscale Modeling and Simulation of Composite Materials and Structures

Multiscale Methods in Computational Mechanics. *In:* de Borst, R. & Ramm, E. (eds.). Springer Berlin / Heidelberg.

Fish, J. & Shek, K. (1999) Finite deformation plasticity for composite structures: Computational models and adaptive strategies. *Computer Methods in Applied Mechanics and Engineering,* 172(1–4)**,** 145-174.

Gao, K., Fu, S., Gibson Jr, R. L., Chung, E. T. & Efendiev, Y. (2014) Generalized Multiscale Finite-Element Method (GMsFEM) for elastic wave propagation in heterogeneous, anisotropic media.

Ghosh, S., Bai, J. & Raghavan, P. (2007) Concurrent multi-level model for damage evolution in microstructurally debonding composites. *Mechanics of Materials,* 39(3)**,** 241-266.

Ghosh, S. & Paquet, D. (2013) Adaptive concurrent multi-level model for multi-scale analysis of ductile fracture in heterogeneous aluminum alloys. *Mechanics of Materials,* 65(0)**,** 12-34.

Gitman, I. M., Askes, H. & Sluys, L. J. (2007) Representative volume: Existence and size determination. *Engineering Fracture Mechanics,* 74(16)**,** 2518-2534.

Gitman, I. M., Askes, H. & Sluys, L. J. (2008) Coupled-volume multi-scale modelling of quasi-brittle material. *European Journal of Mechanics - A/Solids,* 27(3)**,** 302-327.

González, C. & Llorca, J. (2006) Multiscale modeling of fracture in fiber-reinforced composites. *Acta Materialia,* 54(16)**,** 4171-4181.

González, C. & Llorca, J. (2007) Mechanical behavior of unidirectional fiber-reinforced polymers under transverse compression: Microscopic mechanisms and modeling. *Composites Science and Technology,* 67(13)**,** 2795-2806.

Graham, S. & Yang, N. (2002) Representative volumes of materials based on microstructural statistics.

Greco, F., Leonetti, L. & Lonetti, P. (2013) A two-scale failure analysis of composite materials in presence of fiber/matrix crack initiation and propagation. *Composite Structures,* 95(0)**,** 582-597.

Hashin, Z. (1965) On elastic behaviour of fibre reinforced materials of arbitrary transverse phase geometry. *Journal of the Mechanics and Physics of Solids,* 13(3)**,** 119-134.

Hautefeuille, M., Colliat, J. B., Ibrahimbegovic, A., Matthies, H. G. & Villon, P. (2012) A multi-scale approach to model localized failure with softening. *Computers & Structures,* 94-95**,** 83-95.

Hill, R. (1963) Elastic properties of reinforced solids: Some theoretical principles. *Journal of the Mechanics and Physics of Solids,* 11(5)**,** 357-372.

Hillerborg, A., Modéer, M. & Petersson, P. E. (1976) Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements. *Cement and Concrete Research,* 6(6)**,** 773-781.

Hou, T. Y. & Wu, X.-H. (1997) A Multiscale Finite Element Method for Elliptic Problems in Composite Materials and Porous Media. *Journal of Computational Physics,* 134(1)**,** 169-189.

Huang, Y., Yang, Z., Ren, W., Liu, G. & Zhang, C. (2015) 3D meso-scale fracture modelling and validation of concrete based on in-situ X-ray Computed Tomography images using damage plasticity model. *International Journal of Solids and Structures,* 67–68**,** 340-352.

Kaczmarczyk, Ł., Pearce, C. J., Bićanić, N. & de Souza Neto, E. (2010) Numerical multiscale solution strategy for fracturing heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering,* 199(17–20)**,** 1100-1113.

Kanouté, P., Boso, D., Chaboche, J. & Schrefler, B. (2009) Multiscale Methods for Composites: A Review. *Archives of Computational Methods in Engineering,* 16(1)**,** 31-75.

Kassner, M. E., Nemat-Nasser, S., Suo, Z., Bao, G., Barbour, J. C., Brinson, L. C., Espinosa, H., Gao, H., Granick, S., Gumbsch, P., Kim, K.-S., Knauss, W., Kubin, L., Langer, J., Larson, B. C., Mahadevan, L., Majumdar, A., Torquato, S. & van Swol, F. (2005) New directions in mechanics. *Mechanics of Materials,* 37(2–3)**,** 231-259.

Li, Y., McDowell, D. L. & Zhou, M. (2013) Computational prediction of fracture toughness of polycrystalline metals. *13th International Conference on Fracture, Beijing, China,* (June 16–21, 2013).

Loehnert, S. & Belytschko, T. (2007) A multiscale projection method for macro/microcrack simulations. *International Journal for Numerical Methods in Engineering,* 71(12)**,** 1466-1482.

López, C. M., Carol, I. & Aguado, A. (2007) Meso-structural study of concrete fracture using interface elements. II: compression, biaxial and Brazilian test. *Materials and Structures,* 41(3)**,** 601-620.

Markovic, D. & Ibrahimbegovic, A. (2004) On micro–macro interface conditions for micro scale based FEM for inelastic behavior of heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering,* 193(48–51)**,** 5503-5523.

Miehe, C. & Bayreuther, C. G. (2007) On multiscale FE analyses of heterogeneous structures: from homogenization to multigrid solvers. *International Journal for Numerical Methods in Engineering,* 71(10)**,** 1135-1180.

Nguyen, V. P., Lloberas-Valls, O., Stroeven, M. & Sluys, L. J. (2012a) Computational homogenization for multiscale crack modeling. Implementational and computational aspects. *International Journal for Numerical Methods in Engineering,* 89(2)**,** 192-226.

Nguyen, V. P., Stroeven, M. & Sluys, L. J. (2011) Multiscale continuous and discontinuous modeling of heterogeneous materials: a review on recent developments. *Journal of Multiscale Modelling,* 03(04)**,** 229-270.

Nguyen, V. P., Stroeven, M. & Sluys, L. J. (2012b) An enhanced continuous–discontinuous multiscale method for modeling mode-I cohesive failure in random heterogeneous quasi-brittle materials. *Engineering Fracture Mechanics,* 79(0)**,** 78-102.

Oden, J. T., Belytschko, T., Babuska, I. & Hughes, T. J. R. (2003) Research directions in computational mechanics. *Computer Methods in Applied Mechanics and Engineering,* 192(7–8)**,** 913-922.

Oden, J. T., Vemaganti, K., Moës, N. (1999) Hierarchical modeling of heterogeneous solids. Computer Methods in Applied Mechanics and Engineering, 172(1–4), 3–25.

Park, K. & Paulino, G. H. (2012) Computational implementation of the PPR potential-based cohesive model in ABAQUS: Educational perspective. *Engineering Fracture Mechanics,* 93**,** 239-262.

Phu Nguyen, V., Lloberas-Valls, O., Stroeven, M. & Johannes Sluys, L. (2010) On the existence of representative volumes for softening quasi-brittle materials – A failure zone averaging scheme. *Computer Methods in Applied Mechanics and Engineering,* 199(45–48)**,** 3028-3038.

Ren, W., Yang, Z., Sharma, R., Zhang, C. & Withers, P. J. (2015) Two-dimensional X-ray CT image based meso-scale fracture modelling of concrete. *Engineering Fracture Mechanics,* 133**,** 24-39.

Ruiz, G., Pandolfi, A. & Ortiz, M. (2001) Three-dimensional cohesive modeling of dynamic mixed-mode fracture. *International Journal for Numerical Methods in Engineering,* 52(12)**,** 97-120.

Simulia/Abaqus. - Abaqus theory manual. Online version 6.12.

Su, X., Yang, Z. & Liu, G. (2010) Monte Carlo simulation of complex cohesive fracture in random heterogeneous quasi-brittle materials: a 3D study. *International Journal of Solids and Structures*, 47(17), 2336-2345.

Trias, D., Costa, J., Fiedler, B., Hobbiebrunken, T. & Hurtado, J. E. (2006a) A two-scale method for matrix cracking probability in fibre-reinforced composites based on a statistical representative volume element. *Composites Science and Technology,* 66(11–12)**,** 1766-1777.

Trias, D., Costa, J., Mayugo, J. A. & Hurtado, J. E. (2006b) Random models versus periodic models for fibre reinforced composites. *Computational Materials Science,* 38(2)**,** 316-324.

Trias, D., Costa, J., Turon, A. & Hurtado, J. E. (2006c) Determination of the critical size of a statistical representative volume element (SRVE) for carbon reinforced polymers. *Acta Materialia,* 54(13)**,** 3471-3484.

Vaughan, T. J. & McCarthy, C. T. (2011) Micromechanical modelling of the transverse damage behaviour in fibre reinforced composites. *Composites Science and Technology,* 71(3)**,** 388-396.

Wang, X. F., Yang, Z. J., Yates, J. R., Jivkov, A. P. & Zhang, C. (2015) Monte Carlo simulations of mesoscale fracture modelling of concrete with random aggregates and pores. *Construction and Building Materials,* 75**,** 35-45.

Xu, Q., Chen, J., Li, J. & Wang, M. (2013) Multi-scale numerical model for simulating concrete material based on fractal theory. *Acta Mechanica Solida Sinica,* 26(4)**,** 344-352.

Xu, X. F. & Chen, X. (2009) Stochastic homogenization of random elastic multi-phase composites and size quantification of representative volume element. *Mechanics of Materials,* 41(2)**,** 174-186.

Xu, X. F. & Graham-Brady, L. (2005) A stochastic computational method for evaluation of global and local behavior of random elastic media. *Computer Methods in Applied Mechanics and Engineering,* 194(42–44)**,** 4362-4385.

Yang, Z. J., Su, X. T., Chen, J. F. & Liu, G. H. (2009) Monte Carlo simulation of complex cohesive fracture in random heterogeneous quasi-brittle materials. *International Journal of Solids and Structures,* 46(17)**,** 3222-3234.

Yu, R. C., Zhang, X. & Ruiz, G. (2008) Cohesive modeling of dynamic fracture in reinforced concrete. *Computers and Concrete,* Volume 5, Issue 4, 2008, pp.389-400.

NOMENCLATURE

Acronyms:

MsSFrM	-multiscale stochastic fracture mechanics.
MeE	-meso-scale element.
MaE	-macro-scale element.
CIE	-cohesive interface element.
BCs	-boundary conditions.
MSA	-merging and splitting approach.
CDA	-crack path decomposition approach.

Boundary conditions:

B1-X, B2-X	-meso-scale boundary conditions on X direction.
B1-Y, B2-Y	-meso-scale boundary conditions on Y direction.
G1-X, G2-X	-global-scale boundary conditions used for validation on X direction.
G1-Y, G2-Y	-global-scale boundary conditions used for validation on Y direction.

Algebra:

C	-kernel crack path.

H_k	-hidden key-points rendered on the stochastic crack path C.
$P(C)$	-probability of crack path C evaluated as a surface/path integral.
$P(E\|C)$	-prior probability of neighboring crack E to match the kernel crack path C. This probability is evaluated as a surface/ path integral.
$P(C\|E)$	-Bayesian updated probability of crack C given an overlapped neighboring crack path E. This probability is evaluated using a Bayesian inference model.
C_m	-set of crack paths over a neighboring cluster.
L_i	-reduced order integration line fitted on a random crack path.
σ_f	-ultimate strength equivalent to the onset of fracture.
$\sigma_{f,u} = \sigma(u)$	-softening evolution based on the ultimate strength pitch.
G_f	-fracture energy
$G_{f,short}$	-fracture energy evaluated for short overlapping criteria.
$G_{f,long}$	-fracture energy evaluated for long overlapping criteria.
u	-finite displacement.
u_f	-final displacement.
d	-scalar degradation variable.

d_{short} -scalar degradation variable evaluated for short overlapping criteria.

d_{long} -scalar degradation variable evaluated for long overlapping criteria.

V_f -fibre volume fraction.

V_m -matrix volume fraction.

www.elsevier.com/locate/ijpe

Non-stationary Demand Forecasting by Cross-Sectional Aggregation

Bahman Rostami-Tabar, Mohamed Zied Babai, Yves Ducq, Aris Syntetos

PII: S0925-5273(15)00381-3
DOI: http://dx.doi.org/10.1016/j.ijpe.2015.10.001
Reference: PROECO6250

To appear in: *Intern. Journal of Production Economics*

Received date: 3 July 2013
Revised date: 23 September 2015
Accepted date: 2 October 2015

Cite this article as: Bahman Rostami-Tabar, Mohamed Zied Babai, Yves Ducc and Aris Syntetos, Non-stationary Demand Forecasting by Cross-Sectiona Aggregation, *Intern. Journal of Production Economics* http://dx.doi.org/10.1016/j.ijpe.2015.10.001

Non-stationary Demand Forecasting by Cross-Sectional Aggregation

Bahman Rostami-Tabar[a], Mohamed Zied Babai[b,*], Yves Ducq[c], Aris Syntetos[d]

[a] *Coventry University, School of Strategy and Leadership, Priory Street, Coventry, CV1 5FB, UK*

[b] *Kedge Business School, 680 cours de la libération, 33400 Talence, France*

[c] *Univ. Bordeaux, IMS, UMR 5218, 33400 Talence, France*

[d] *Cardiff Business School, Cardiff University, Aberconway Building, Cardiff CF10 3EU, UK*

Abstract:

In this paper the relative effectiveness of top-down (TD) versus bottom-up (BU) approaches is compared for cross-sectionally forecasting aggregate and sub-aggregate demand. We assume that the sub-aggregate demand follows a non-stationary Integrated Moving Average (IMA) process of order one and a Single Exponential Smoothing (SES) procedure is used to extrapolate future requirements. Such demand processes are often encountered in practice and SES is one of the standard estimators used in industry (in addition to being the optimal estimator for an IMA process). Theoretical variances of forecast error are derived for the BU and TD approach in order to contrast the relevant forecasting performances. The theoretical analysis is supported by an extensive numerical investigation at both the aggregate and sub-aggregate level, in addition to empirically validating our findings on a real dataset from a European superstore. The results demonstrate the increased benefit resulting from cross-sectional forecasting in a non-stationary environment than in a stationary one. Valuable insights are offered to demand planners and the paper closes with an agenda for further research in this area.

Keywords: Demand Forecasting; Cross-sectional aggregation; Non-Stationary Processes; Single Exponential Smoothing

* Corresponding author: Mohamed Zied BABAI; e-mail: mohamed-zied.babai@kedgebs.com

E-mail addresses: ac1055@coventry.ac.uk (Bahman ROSTAMI-TABAR), SyntetosA@cardiff.ac.uk (Aris A. SYNTETOS), yves.ducq@ims-bordeaux.fr (Yves DUCQ).

1. INTRODUCTION

Demand forecasting is the starting point for most planning and control organizational activities. A considerable part of the forecasting literature has been dedicated to strategies and methods for single time series, but in reality there are often many related time series that can be organized hierarchically and aggregated at several different levels in groups based on products, customers, geography or other features. The hierarchical level at which forecasting is performed then it will depend on the function the forecasts are fed into. With regards to products (or Stock Keeping Units - SKUs) in particular, forecasting at the individual SKU level is required for inventory control, product family forecasts may be required for Master Production Scheduling, forecasts across a group of items ordered from the same supplier may be required for the purpose of consolidating orders, forecasts across the items sold to a specific large customer may determine transportation and routing decisions etc.

One intuitively appealing approach to obtain higher level forecasts is by *cross-sectional (also referred to as hierarchical) aggregation*, which involves aggregating different items (i.e. aggregating the requirements for different items usually in one specific time period) to reduce variability. Existing approaches to cross-sectional forecasting usually involve either a bottom-up (BU) or a top-down (TD) approach (or a combination of the two). When forecasting at the aggregate level is of interest, the former involves the aggregation of individual SKU forecasts to the group level whereas the latter relates to forecasting directly at the group level (i.e. first aggregate requirements and then extrapolate directly at the aggregate level). When the emphasis is on forecasting at the sub-aggregate level, then bottom-up relates to direct extrapolation at the sub-aggregate level whereas top-down involves the disaggregation of the forecasts produced directly at the group level (Gross and Sohl, 1990; Widiarta et al., 2007). An important issue that has attracted the attention of many researchers as well as practitioners over the last few decades is the (relative) effectiveness of such cross-sectional forecasting approaches.

TD and BU are extremely useful towards improving the accuracy of forecasts and plans when leveraged within an S&OP (Sales and Operations Planning) process (Lapide, 2006). The S&OP is a multi-functional process that involves managers from all departments (Sales, Customer Service, Supply Chain, Marketing, Manufacturing, Logistics, Procurement and Finance), where each department requires different levels of demand forecasts (Lapide, 2004). For example, in marketing, forecasting of revenues by product groups and brands is needed, sales departments deal with sales forecasts by customer accounts and/or sales channels,

supply chain managers request SKU level forecasts, while finance requires forecasts that are aggregated into budgetary units in terms of revenues and costs (Bozos and Nikolopoulos, 2011).

In this paper, we study analytically the relative effectiveness of the BU and TD approach when the underlying series follows a non-stationary Integrated Moving Average process of order one, ARIMA(0,1,1), and the forecasting method is the Single Exponential Smoothing (SES) which is the optimal estimator for the process under consideration (Box et al., 2008). Both assumptions bear a significant degree of realism. There is evidence to support the fact that demand often follows non-stationary processes (please refer also to subsection 2.1). Moreover, SES is a very popular forecasting method in industry (Acar and Gardner, 2012; Gardner, 1990, 2006; Taylor, 2003). In terms of the practical relevance of our research we refer to a set of SKUs where a large proportion of them follow an ARIMA(0,1,1) process; this is not an untypical scenario as demonstrated by analysis of empirical datasets including our own empirical investigation.

The question is whether it is appropriate to use sub-aggregate data or one should rather aggregate data to derive the individual and aggregate forecasts. In addition, we analyse the case of non-stationary processes to reveal whether there is an increased benefit resulting from cross-sectional forecasting when departing from the stationarity assumption. To do so we compare the variance of the forecast error obtained based on the aggregate demand (V_{TD}) to that of the sub-aggregate demand (V_{BU}). Comparisons are performed at both the aggregate and sub-aggregate level, in the former case using both a theoretical and a numerical analysis while in the latter case only by means of a numerical simulation (since the mathematical results in that case are intractable). Our analysis is consistent with the fact that companies are often using both levels of forecasting to support different decision-making processes. In addition, it renders the comparison between the two approaches a more fair exercise since one might expect that BU provides more accurate forecasts at the sub-aggregate level and TD works better at the aggregate level (Zotteri et al., 2005).

We mathematically show that the ratio of the variance of forecast error of the top-down to that of the bottom-up approach is equal to one for identical process parameters when compared at the aggregate level. The mathematical analysis is complemented by a numerical experiment to evaluate in detail the conditions under which one approach outperforms the other. Such an experiment also allows the introduction of non-identical process parameters of the sub-aggregate series (a condition that cannot be considered mathematically) and the comparison at the sub-aggregate level. In addition, an empirical investigation is also

conducted to assess the validity of the results on real data from a European superstore. Important managerial insights are derived based on the above research and tangible suggestions are offered to practitioners dealing with inventory forecasting problems.

To the best of our knowledge, the only papers directly relevant to our work are those by Widiarta et al. (2007, 2008, 2009) and Sbrana and Silvestrini (2013). The researchers evaluated analytically at the aggregate level the effectiveness of the TD and BU approaches under the assumption of AR(1) (Auto-Regressive process of order 1), MA(1) and IMA(1,1) processes. Our additional contribution to the literature is threefold: (i) we analyse the superiority conditions of BU and TD approaches at both the aggregate and the sub-aggregate levels of forecasting by means of both analytical and simulation work, (ii) through a more detailed sensitivity analysis using simulation, we investigate the impact of all the process and control parameters on the comparative performance of the two approaches, and (iii) we analyse and validate empirically our theoretical results on real data noting that none of the previous theoretical work comparing BU and TD approaches has been validated empirically.

With regards to this last point, it is important to note that rather recently Athanasopoulos et al. (2009) and Hyndman et al. (2011) have proposed a new approach to handling hierarchical time series forecasting. This approach does not emphasise the estimator being used to extrapolate requirements (i.e. any forecasting method may actually be used) but rather the weighted contribution of the forecasts produced at all levels of a given hierarchy (to appropriately retain important information that may be available at any hierarchical level) for the purpose of producing a required forecast at a particular level. Despite the fact that this approach lacks analytical insights it has been shown to perform well in practice and thus it is further considered (in addition to the BU and TD approaches) in the empirical part of our investigation.

The remainder of our paper is structured as follows. In Section 2 we provide a review of the literature on demand aggregation related issues. In Section 3 we describe the assumptions and notations used in this study and we conduct an analytical evaluation of the variance of the forecast error related to both the BU and TD approaches, followed by a simulation study performed in Section 4. We conduct an empirical investigation in Section 5 and the paper concludes in Section 6 with the implications of our work for real world practices along with an agenda for further research in this area.

2. LITERATURE REVIEW

Demand forecasting for sales and operations management often concerns many items, perhaps hundreds of thousands, simultaneously. The conventional forecasting approach is to extrapolate the data series for each SKU individually. However, most businesses have natural groupings of SKUs; that is, the SKUs may be aggregated to get higher levels of forecasts across different dimensions such as product families, geographical areas, customer types, supplier types etc. (Chen and Boylan, 2007). Such an approach enables the potential identification of time series components such as trend or seasonality that may be hidden or not particularly prevalent at the individual SKU level. Group approaches for example are known to offer considerable benefits towards the estimation of seasonal indices (Chen and Boylan 2008). Most of the forecasting literature in this area has looked at the comparative performance of the top-down (TD) and the bottom-up (BU) approach. The findings with regards to the performance of these approaches are mixed.

Many researchers have provided evidence in favour of the TD approach. Gross and Sohl (1990) for example, numerically found that the TD approach (in conjunction with an appropriate disaggregation method) provided better estimates than BU forecasting in two out of three product lines examined. Fliedner (1999) evaluated by means of simulation the forecast system performance at the aggregate level resulting from varying degrees of cross correlation between two sub-aggregate time series. The sub-aggregate items were assumed to follow a Moving Average process of order one, MA(1), and the forecasting methods considered were SES and the Simple Moving Average (SMA). This research showed the forecast performance at the aggregate level to benefit from the TD approach. Barnea and Lakonishok (1980) examined the effectiveness of BU and TD on forecasting corporate performance. They reported that positive cross-correlation contributes to the superiority of forecasts based on aggregate data (TD).

On the other hand, Orcutt et al. (1968) and Edwards and Orcutt (1969) argued that information loss is substantial when aggregating and therefore the bottom-up approach provides more accurate forecasts. Dangerfield and Morris (1992) and Gordon et al. (1997) used a subset of the M-competition[b] data (Makridakis et al., 1982) to examine the performance of TD and BU approaches on sub-aggregate demand forecasting. They found that forecasts by the BU approach were more accurate in most situations especially when

[b] The M Competition is an empirical forecast accuracy comparison exercise introduced by Prof. Makridakis.

items were highly correlated or when one item dominated the aggregate series. Weatherford et al. (2001) evaluated the performance of BU and TD approaches to obtain the required forecasts for hotel revenue management. The data they considered was perceived as very typical within the hotel industry. They experimented with four different approaches (fully disaggregated, aggregating by rate category only, aggregating by length of stay only, and aggregating by both rate category (i.e. the price per night) and length of stay) to get detailed forecasts by day of arrival, duration of stay and rate category. The results of their study showed that a purely sub-aggregate forecast strongly outperformed even the best aggregate forecast.

Some authors take a contingent approach and analyse the conditions under which one method produces more accurate forecasts than the other. Shlifer and Wolff (1979) evaluated analytically the superiority of BU and TD on forecasting sales for specific and entire market segments. They specified the conditions under which BU is preferred to TD and vice versa. Such superiority was found to be a function of the number of markets, market size and forecast horizon. They mentioned that BU is preferable for the purpose of forecasting the aggregate series based on their observations in real situations. However, when the comparison was performed at the sub-aggregate level, they found that BU performs better for small marker segments, while both BU and TD perform equally well for large segments. Lütkepohl (1984) evaluated the performance of BU and TD approaches for forecasting at the aggregate level by using the mean squared error. It was shown that it might be preferable to forecast aggregate time series using a BU strategy when the data generation process is known. However, if the ARIMA processes used for forecasting, including the order of process and parameters, were estimated from a given set of time series data then the TD approach outperformed BU. Widiarta et al. (2007) studied analytically the conditions under which one approach outperforms the other for forecasting the item level demands when the sub-aggregate items follow a first-order autoregressive [AR(1)] process with the same autoregressive parameter for all the items and when SES is used to extrapolate future demand requirements. They found that the superiority of each approach is a function of the autoregressive parameter. Widiarta et al. (2008, 2009) also evaluated analytically the effectiveness of TD and BU approaches at the sub-aggregate and aggregate level. They showed that when all sub-aggregate items follow an MA(1) process with identical moving average parameters, there is no difference in the relative performance of TD and BU forecasting as long as the optimal smoothing constant is used in both approaches. Subsequently, they conducted a simulation analysis considering non-identical process

parameters for sub-aggregate items and concluded that there is significant difference between the two approaches. The superiority of each approach was a function of the moving average parameter, the cross-correlation and the proportion of a sub-aggregate component's contribution to the aggregate demand. Sbrana and Silvestrini (2013) evaluated the effectiveness of BU and TD approaches when forecasting the aggregate demand using a multivariate exponential smoothing framework. They established the necessary and sufficient condition for the equality of mean squared errors (MSEs) of the two approaches. In addition, they showed that the relative forecasting accuracy of TD and BU depends on the parametric structure of the underlying framework.

In summary, both BU and TD approaches appear to be associated with more accurate forecasts depending on the level of comparison, structure of the series and cross-correlation related assumptions. It is easy to observe that most of the literature dealing with the issue of aggregation for forecasting purposes focuses only on stationary series and it does not consider the most realistic case of non-stationary processes. There is considerable evidence to suggest that inventory demand is non-stationary and thus relevant processes could be eventually assumed for representing their underlying structure. In the following subsection we provide an overview of the literature on the validity of the nonstationary assumption for real world applications.

Before we close this section it should be noted that recently Athanasopoulos et al. (2009) and Hyndman et al. (2011) have proposed a new approach, referred to as 'the optimal method', to handling hierarchical time series forecasting. As discussed in the previous section, the approach under concern is based on independently forecasting the series at all levels of a given hierarchy and then using a regression model to optimally combine and reconcile these forecasts at a particular desired forecast output level. By means of a simulation study using ARIMA type series and an empirical investigation using Australian tourism demand data, Hyndman et al.(2011) have shown that the optimal method outperforms both the TD and the BU approaches and as such we further consider it as a benchmark in the empirical part of our work.

2.1. The validity of the non-stationary demand assumption

Compared to stationary demand processes, nonstationary processes have received less attention in the academic literature (Bijvank and Vis, 2011) although there is evidence that most of the forecasting and inventory control problems occur in situations where demand is

nonstationary (and partially observed) (Treharne and Sox, 2002). Naturally this may be attributed to the fact that the nonstationarity assumption complicates the relevant analyses and limits the theoretical results that may be obtained making it very difficult to determine an optimal forecast and stocking levels (Shang, 2012).

Nonstationary demand is the rule rather than the exception in most industries nowadays. The nonstationarity may arise due to many reasons such as: (1) product life cycles with multi-stages, (2) technological innovation and reduced product life, (3) seasonal effects, (4) volatile customer preferences, (5) changes in economic conditions, (6) exchange rate fluctuations, etc. (Li et al., 2011). Companies in all markets are introducing new products at a higher frequency with increasingly shorter life cycles. For instance, in the high-tech industry, the products have relatively short life cycles and their demand patterns are generally considered as nonstationary (Chien et al., 2008; Graves and Willems, 2000, 2008; Raghunathan, 2001).

Furthermore, nonstationary demand processes have been observed in the wholesaling and retailing industry. Martel et al. (1995) argued that in the grocery distribution, because of the various promotion mechanisms such as weekly special promotions, national television advertising campaigns, etc., demand gets clearly nonstationary. Erkip et al. (1990) and Lee et al. (1997) empirically found that demands of consumer products are nonstationary and highly autocorrelated. Lee et al. (2000) used panel data to examine the weekly sales patterns of 165 SKUs at a supermarket. They found that 150 out of the 165 SKUs analysed demonstrated nonstationary behaviour with high autocorrelation. Ali et al. (2011) experimented with a demand dataset of 1798 SKUs from a major European supermarket in Germany. They found that around 30% of the SKUs follow a nonstationary process and further an 80% of them follow an ARIMA(0,1,1) process. Moreover, Mitchell and Niederhausen (2010) noted that the nature of a nonstationary demand processes is consistent with the nature of retail demand for a wide variety of merchandise including apparel, consumer electronics, toys and other holiday items, patio furniture and other summer seasonal merchandise and school supplies.

Another sector where it was reported that demand follows a nonstationary process is the tourism (Goh and Law, 2002). Finally, Tunc et al. (2011) confirmed that nonstationary stochastic demands are very common in all industrial settings associated with seasonal patterns, trends, business cycles, and limited-life items.

There is also evidence that demand may follow an ARIMA(0,1,1) process in particular (which is the process considered in this study). This process has often been found to be useful in inventory control problems and econometrics (Box et al., 2008). More generally, Mahajan

and Desai (2011) argued that retailers often face a nonstationary demand that follows an ARIMA(0,1,1) process.

In this study we compare the performance of BU and TD approaches on demand forecasting under the assumption of a nonstationary ARIMA(0,1,1) process. In the next section we analyse theoretically the forecasting effectiveness of these approaches.

3. THEORETICAL ANALYSIS

In this section we derive the variance of the forecast error associated with the TD and BU approaches. These approaches work as follows:

The top-down approach consists of the following steps: i) sub-aggregate demand items are aggregated; ii) the forecast of aggregate demand is produced by applying SES at the aggregate level, and iii) the forecast is disaggregated back to the original level by applying an appropriate disaggregation method, if a sub-aggregate forecast is needed. Various proportional approaches may be used to disaggregate the TD forecasts. The reader is referred to Gross and Sohl (1990) for more details about such approaches.

In the bottom-up approach: i) sub-aggregate demand forecasts are produced directly for the sub-aggregate items; ii) the aggregate forecast (if needed) is obtained by combining individual forecasts for each SKU, i.e. potentially a separate forecasting model is used for each item in the product family (Zotteri et al., 2005). These approaches are presented schematically in Figure 1. The presentation style follows that adopted by Mohammadipour et al. (2012).

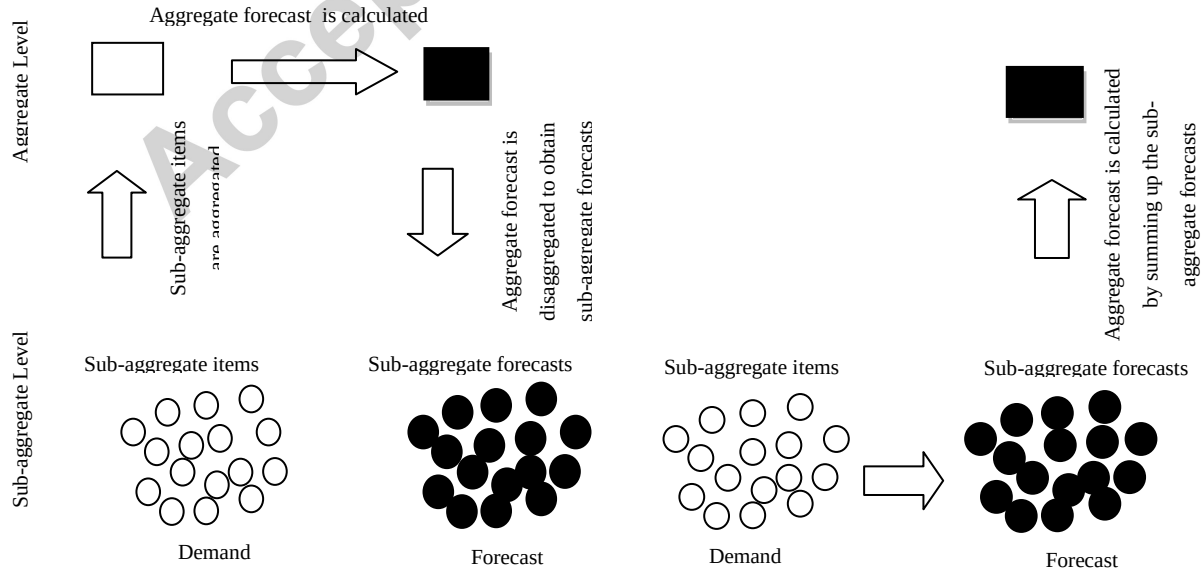


Figure 1. Graphical representation of the TD (left) and BU (right) approach

Comparisons may be performed at both the aggregate and sub-aggregate level although in our theoretical analysis the comparisons are performed only at the former level since (analytical) results regarding the latter are intractable. However, in the simulation study that follows the theoretical analysis we relax various assumptions and we also present results for both levels of comparison.

3.1. Notation and Assumptions

We denote by:

$d_{i,t}$: Sub-aggregate item demand *i* in period *t*

$\rho_{i,j}$: Correlation between the error term of sub-aggregate item *i* and *j* (cross-correlation)

D_t: Aggregate demand in period *t*

$\varepsilon_{i,t-k}$: Independent random variable for sub-aggregate item demand *i* in period *t*, normally distributed with zero mean and variance σ^2

ε'_{t-k}: Independent random variable for aggregate demand in period *t*, normally distributed with zero mean and variance σ'^2

$f_{i,t}$: Forecast of sub-aggregate demand in period *t*, the forecast produced in *t-1* for the demand in *t*.

F_t : Forecast of aggregate demand in period *t*, the forecast produce in *t*-1 for the demand in *t*.

α_i: Smoothing constant used in the Single Exponential Smoothing method for each sub-aggregate item *i* in the BU approach, $0 \le \alpha_i \le 1$

α_{TD}: Smoothing constant used in the Single Exponential Smoothing method for aggregate demand in TD approach, $0 \le \alpha_{\mathrm{TD}} \le 1$

p_i: the relative weight of sub-aggregate item *i*'s contribution to the aggregate family, where $\sum_{i=1}^{N} p_i = 1$.

V_{BU}: Variance of Forecast Error of the BU approach

V_{TD}: Variance of Forecast Error of the TD approach

V_{OP}: Variance of Forecast Error of the optimal method

θ_i : Moving average parameter of sub-aggregate item demand *i*, $|\theta_i| < 1$

θ' : Moving average parameter of aggregate demand, $|\theta'| < 1$

C_i : Constant value of sub-aggregate item demand *i*

C' : Constant value of aggregate demand.

N: the total number of sub-aggregate items.

We assume that all the sub-aggregate demand series $d_{i,t}$ follow an Integrated Moving Average process of order one, ARIMA(0,1,1), that can be mathematically written by (1).

$$\begin{cases} d_{i,1} = C_i + d_{i,0} & i = 1,2,\ldots,N \\ d_{i,t} = d_{i,t-1} + \varepsilon_{i,t} - \theta_i \varepsilon_{i,t-1} & i = 1,2,\ldots,N;\ t = 2,3,\ldots \end{cases} \quad (1)$$

From (1) it is obvious that the demand in the next period is the demand in the current period plus an error term. By writing and expanding (1) in a recursive form we have:

$$d_{i,t} = \left(C_i + d_{i,0}\right) + \varepsilon_{i,t} + \alpha_i \varepsilon_{i,t-1} + \alpha_i \varepsilon_{i,t-2} + \ldots + \alpha_i \varepsilon_{i,1} \quad (2)$$

where $\alpha_i = 1 - \theta_i$. We note that only under this condition on α_i, SES is optimal as it provides the minimum mean square error forecasts for the ARIMA(0,1,1) process. Here we consider the smoothing constant values as a control parameter determined by forecasters which varies between 0 and 1. Obviously, under this condition since $0 \le \alpha_i \le 1$, θ_i will only take values between 0 and 1 and does not cover the whole range of $-1 \le \theta_i \le 1$.

However, the theoretical analysis is still valid for the whole range of $-1 \le \theta_i \le 1$. In addition, in the simulation analysis we will relax this assumption to cover the whole range of $-1 \le \theta_i \le 1$ when the value of the smoothing constant is fixed.

We assume that all the sub-aggregate demand process parameters are identical ($\theta_1 = \theta_2 = \theta_3 = \ldots = \theta_N$). This assumption is considered only for the purpose of the theoretical analysis and, as above, it is also relaxed in the simulation part of our work. The assumption under concern implies that the aggregate demand also follows an ARIMA (0,1,1) process. If $\theta_1 \neq \theta_2 \neq \theta_3 \neq \ldots \neq \theta_N$ then the sum of the sub-aggregate items is not necessarily an ARIMA(0,1,1) process (Granger and Morris, 1976).

The aggregate demand in period t, D_t can be expressed as the sum of the demands of the sub-aggregate items, i.e. $D_t = \sum_{i=1}^{N} d_{i,t}$.

The forecasting method considered in this study is the Single Exponential Smoothing (SES); this method is being applied in very many companies and in particular in an inventory production planning environment due to its simplicity (Gardner, 1990). At this point we should mention that the MSE is formally defined as $MSE = Var(d_t - f_t) + Bias^2$, where *Bias* is defined as the expected forecast error and equals to $E(d_t - f_t)$ (Syntetos, 2001).

Note that under the condition that $\alpha_i = 1 - \theta_i$, SES is an unbiased estimator of the demand process considered in this study. Therefore *Bias* = 0 which means that the variance of the forecast error is equal to the mean square error. Using *SES*, the forecast of sub-aggregate demand *i* in period t produced at the end of period *t*-1 is

$$f_{i,t} = (C_i + d_{i,0}) + \alpha_i \varepsilon_{i,t-1} + \alpha_i \varepsilon_{i,t-2} + \ldots + \alpha_i \varepsilon_{i,1} \tag{3}$$

We further assume that the standard deviation of the error term in (1) is significantly smaller than the expected value of the demand, so should demand be generated the probability of a negative value is negligible.

3.2. Comparison of the Variance of Forecast Error

We calculate the ratio of the variance of the forecast error corresponding to the TD approach (V_{TD}) to the variance of the forecast error associated with the BU approach (V_{BU}). A ratio that is lower than one implies a benefit in favour of the TD approach. Conversely, if the ratio is greater than one then the BU approach performs better (and if the ratio is equal to one both strategies perform the same).

We begin the analysis by deriving the $V_{BU,}$ which is defined as follows:

$$V_{BU} = Var\left(D_t - \sum_{i=1}^{N} f_{i,t}\right) = Var\left(\sum_{i=1}^{N} d_{i,t} - \sum_{i=1}^{N} f_{i,t}\right) = Var\left(\sum_{i=1}^{N} (d_{i,t} - f_{i,t})\right) \tag{4}$$

by substituting (2) and (3) in (4) we have:

$$V_{BU} = Var\left(\sum_{i=1}^{N} \varepsilon_{i,t}\right) \tag{5}$$

Since $Var(\varepsilon_{i,t}) = \sigma_i^2$ and $Cov(\varepsilon_{i,t}, \varepsilon_{j,t}) = \rho_{i,j}\sigma_i\sigma_j$ we have:

$$V_{BU} = \sum_{i=1}^{N} \sigma_i^2 + 2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \rho_{i,j}\sigma_i\sigma_j \tag{6}$$

We now derive the variance of the forecast error for the TD approach. As discussed above, it has been shown that when the sub-aggregate items follow an IMA (1,1) process, the aggregate family demand also follows an IMA (1,1) process (Granger and Morris, 1976). The family aggregate process is defined as follows:

$$\begin{cases} D_1 = C' + D_0 \\ D_t = D_{t-1} + \varepsilon'_t - (1-\alpha_{TD})\varepsilon'_{t-1} \quad t = 2,3,... \end{cases} \tag{7}$$

where $\theta' = 1-\alpha_{TD}$.

Considering $\theta_1 = \theta_2 = \theta_3 = \ldots = \theta_N = \theta$ results in the same theta also in the aggregate demand so, $\theta' = \theta$. Now by considering $\theta' = 1-\alpha_{TD}$ and $\theta = \theta'$, it is obvious that the optimal smoothing constant for the aggregate demand is $\alpha_{TD} = 1-\theta$, which is equal to the optimal smoothing constant for the sub-aggregate process.

The aggregate demand and its forecast can be expressed as a function of the error terms, so we have:

$$D_t = (C' + D_0) + \varepsilon'_t + \alpha_{TD}\varepsilon'_{t-1} + \alpha_{TD}\varepsilon'_{t-2} + \ldots + \alpha_{TD}\varepsilon'_1 \tag{8}$$

Knowing that $\varepsilon'_t = \sum_{i=1}^{N} \varepsilon_{i,t}$, we obtain

(9)

$$Var(\varepsilon'_t) = \sum_{i=1}^{N} Var(\varepsilon_{i,t}) + 2\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} Cov(\varepsilon_{i,t}, \varepsilon_{j,t})$$

The aggregate forecast is

$$F_t = (C' + D_0) + \alpha_{TD}\varepsilon'_{t-1} + \alpha_{TD}\varepsilon'_{t-2} + \ldots + \alpha_{TD}\varepsilon'_1 \tag{10}$$

The variance of the TD forecast error is defined as:

$$V_{TD} = Var\left[D_t - F_t\right] \tag{11}$$

By substituting (8) and (10) into (11), we have:

$$V_{TD} = Var\left[\varepsilon'_t\right] \tag{12}$$

By substituting (9) into (12) we have:

$$V_{TD} = \sum_{i=1}^{N}\sigma_i^2 + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\rho_{i,j}\sigma_i\sigma_j \tag{13}$$

Proposition. If all the sub-aggregate demand items follow an ARIMA(0,1,1) process with identical moving average parameters ($\theta_1 = \theta_2 = \theta_3 = \ldots = \theta_N$) and the optimal smoothing constant value is used to forecast both the sub-aggregate and aggregate demand, then the performance of the TD and BU approaches for forecasting aggregate demand is identical ($V_{TD} = V_{BU}$). The proof of the proposition follows directly from (6) and (13).

This finding is in agreement with the results reported by Widiarta et al. (2009) which theoretically show that there is no significant difference between the TD and BU approaches on forecasting aggregate demand when all sub-aggregate items follow an MA(1) process with identical process parameters.

4. SIMULATION STUDY

In this section, we perform a simulation study to evaluate the relative performance of the TD over the BU approach under some more realistic assumptions. In particular we consider the following scenarios: *i*) a simulation study at the aggregate level for non-identical ($\theta_1 \neq \theta_2 \neq \ldots \neq \theta_N$) process parameters; *ii*) a simulation investigation to discuss the effectiveness of the BU and TD approach compared at the sub-aggregate level for non-identical ($\theta_1 \neq \theta_2 \neq \ldots \neq \theta_N$) process parameters. In both cases, the search procedure has been performed in the whole range of $-1 \leq \theta_i \leq 1$.

When the underlying process follows an ARIMA(0,1,1) representation, as θ_i moves from +1 toward -1 the resulting underlying structure changes considerably. When θ_i is negative, the autocorrelation parameter exhibits a smooth exponential decay with positive values and the autocorrelation spans all time lags (not only lag 1). For example for θ_i=-0.9 the autocorrelation is very close to +1. As we move up towards $\theta_i \cong$+1 the autocorrelation reduces but still remains positive and for high positive values of θ_i it becomes close to zero meaning that the series are random.

By considering many SKUs in the simulation experiment, the presentation of results and the evaluation of the impact of different parameters on the ratio of V_{TD} / V_{BU} becomes complex. Therefore, we restricted the simulation analysis to a family of two SKUs to obtain meaningful insights. This is in concordance with most of the earlier papers using simulation approaches as they have also restricted the number of items to two (Dangerfield and Morris, 1992; Fliedner, 1999; Widiarta et al., 2008, 2009).

The parameter values for our simulation experiment are presented in Table 1.

Table 1. Parameters of the simulation experiment

μ_i	σ_i^2	θ_i	ρ_{ij}	N° Sub-Aggregate	N° Replications	N° Time Periods
400	900	-0.9 : +0.9	-0.9: +0.9	2	100	1000

The sub-aggregate demands in each period are generated randomly subject to the parameters described in Table 1.

The value of σ_i is set to be quite smaller than μ_i to avoid the generation of negative sub-aggregate demand values. Experiments have also been conducted with other values of μ_i and σ_i but these are not reported here as they lead to the same insights.

To generate the demands in each period *t*, we first generate randomly the error terms $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ with a cross-correlation coefficient of ρ_{12} and then we use (1) to generate the correlated sub-aggregate demands. We initialise the generated demand at the value of the mean plus an error term. The simulation experiment has been designed and run in the forecast package in R. For each parameter combination described in Table 1, a demand series of 1000 observations is generated and we introduce 100 replications.

We split the generated demand for each series at both the sub-aggregate and aggregate level, into two parts. The first part (within sample) consists of 700 time periods and is used in

order to fit the model and estimate the parameters. The smoothing constant and the initial value are found based on Maximum Likelihood estimation (Hyndman et al., 2008; Hyndman et al., 2002). Note that for the BU approach, the smoothing constants are optimized for each item individually. Finally, in order to evaluate the performance of the two forecasting approaches, we calculate the value of the variance of the forecast error for the last 300 periods of the simulation (out-of-sample).

The relative benefit of one forecasting approach over the other is measured by V_{TD} / V_{BU}. As previously discussed, a ratio lower than one implies that the TD approach outperforms the BU one whereas a ratio greater than one implies the opposite.

4.1. Comparison at the Aggregate Level

We first analyse the relative performance of the two forecasting approaches at the aggregate level when the sub-aggregate process parameters are not necessarily identical. For each experiment, the ratio of the variance of the forecast error is calculated as $Var(D_t - F_t) \Big/ Var\left(D_t - \sum_{i=1}^{2} f_{i,t}\right)$.

The simulation results show that when the process parameters are identical there is no difference between the BU and the TD approach. Whereas, when the process parameters are not identical, which is more realistic, the results are different. The results for the latter case are presented in Figure 2.

We see that as the cross-correlation coefficient changes from *-0.9* toward *+0.9* the ratio of V_{TD}/V_{BU} is being reduced. The ratio is higher than or equals to 1 when the cross-correlation is negative, equals to zero or takes low positive values. The ratio is lower than 1 only if the cross-correlation is (highly) positive.

The detailed results show that when the moving average parameters, θ_1 and θ_2, take negative values, the performance of BU and TD approaches is always identical regardless of the values of the cross-correlation. One possible explanation for this result is that when the MA parameters of sub-aggregate series take negative values, the optimal value of the smoothing constant is set at the highest value in the considered range which is equal to *0.99* for both approaches. As the smoothing constant for both BU and TD approaches is equal and the same procedure of forecasting is used for the sub-aggregate items and the aggregate one, the aggregate forecast under both BU and TD approaches is the same.

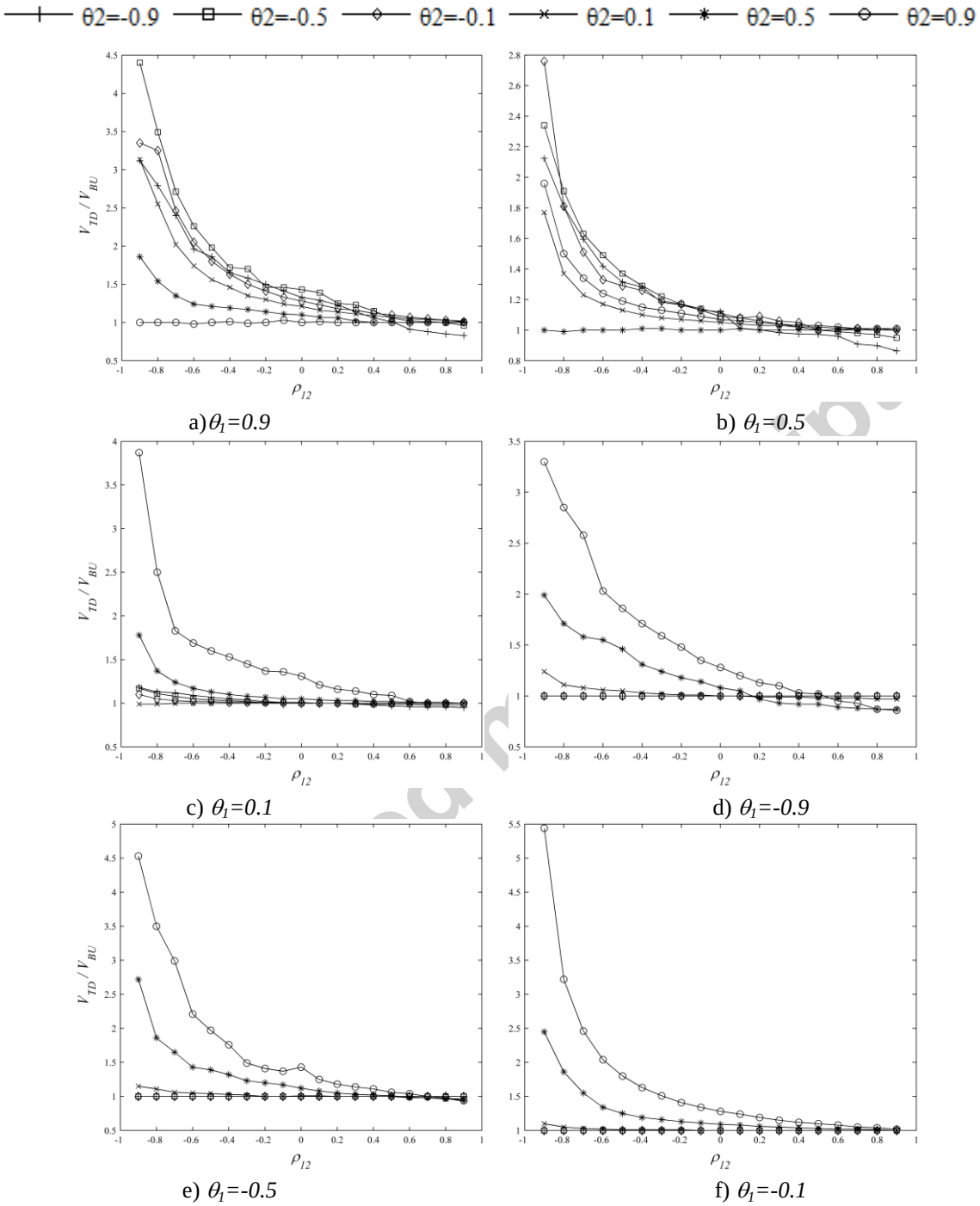


Figure 2. Relative performance of the TD and BU approach in forecasting aggregate demand under different combinations of θ_1, θ_2 and ρ_{12}

When the cross-correlation is positive the superiority of each approach depends on the value and the sign of the moving average parameters, θ_1 and θ_2. The TD approach

outperforms the BU one only when the cross-correlation is (highly) positive and the moving average parameters take high values and have opposite signs, i.e. either $\theta_1<0$ and $\theta_2>0$ or $\theta_1>0$ and $\theta_2<0$.

Note that as the cross-correlation decreases the superiority of the TD approach decreases too. For highly positive cross-correlation, TD outperforms BU with a forecast error variance reduction that can go up to 15%. By decreasing the cross-correlation to 0.5, the maximum benefit of the TD approach decreases to 5% and it tends toward zero when the cross-correlation tends towards zero as well. Under negative cross-correlation, BU outperforms TD.

When the two moving average parameters take opposite signs under the ARIMA(0,1,1) process, this means that one series has positive autocorrelation while the other has a low autocorrelation (series with random fluctuations). In addition, when the cross-correlation is positive there is a tendency for the pair of series to move together in the same direction, so the demand series have the same pattern. When using TD, we sum up all sub-aggregate series to get an aggregate one, so the fluctuations from one series may be cancelled out by those of another resulting in a less random series associated with a lower forecast error.

When the cross-correlation coefficient is negative, for all values of θ_1 and θ_2, with the exception of the case when both are negative, the BU approach performs better. Performance differences are further inflated when the moving average parameters have opposite signs in which case the variance reduction achieved by the BU approach can be as high as 500% for highly negative cross-correlation. For negative cross-correlation, the pair of series moves in the opposite direction, (i.e. if one increases the other decreases), so the sub-aggregate demand series have different patterns of evolution. Combination of different patterns of variation and opposite autocorrelation values lead to a large forecast error for the TD approach and consequently large values of V_{TD} / V_{BU} for very high negative cross-correlation. In these cases it is better to forecast sub-aggregate requirements separately and then aggregate them to get the aggregate forecast.

When the θ_1 and θ_2 values are positive, the ratio is almost equal to 1 for highly positive cross-correlation and greater than 1 for less positive and negative cross-correlation. In the latter case, the ratio of V_{TD} / V_{BU} is increased as θ_1 takes low values and θ_2 is high and vice versa.

In summary, when the sub-aggregate items follow an ARIMA(0,1,1) process and the goal is to forecast at the aggregate demand level, then: *i*) if the autocorrelation of all items is highly positive, the performance of BU and TD is always identical; *ii*) if items have different

autocorrelation patterns, one has a very high positive autocorrelation while the other has an autocorrelation close to zero, the superiority of each approach is affected by the cross-correlation between items; for a highly positive cross-correlation, TD outperforms BU and for a highly negative cross-correlation BU outperforms TD; *iii*) when the autocorrelation for all items is low, BU generally dominates TD, although for highly positive cross-correlation the difference is very low.

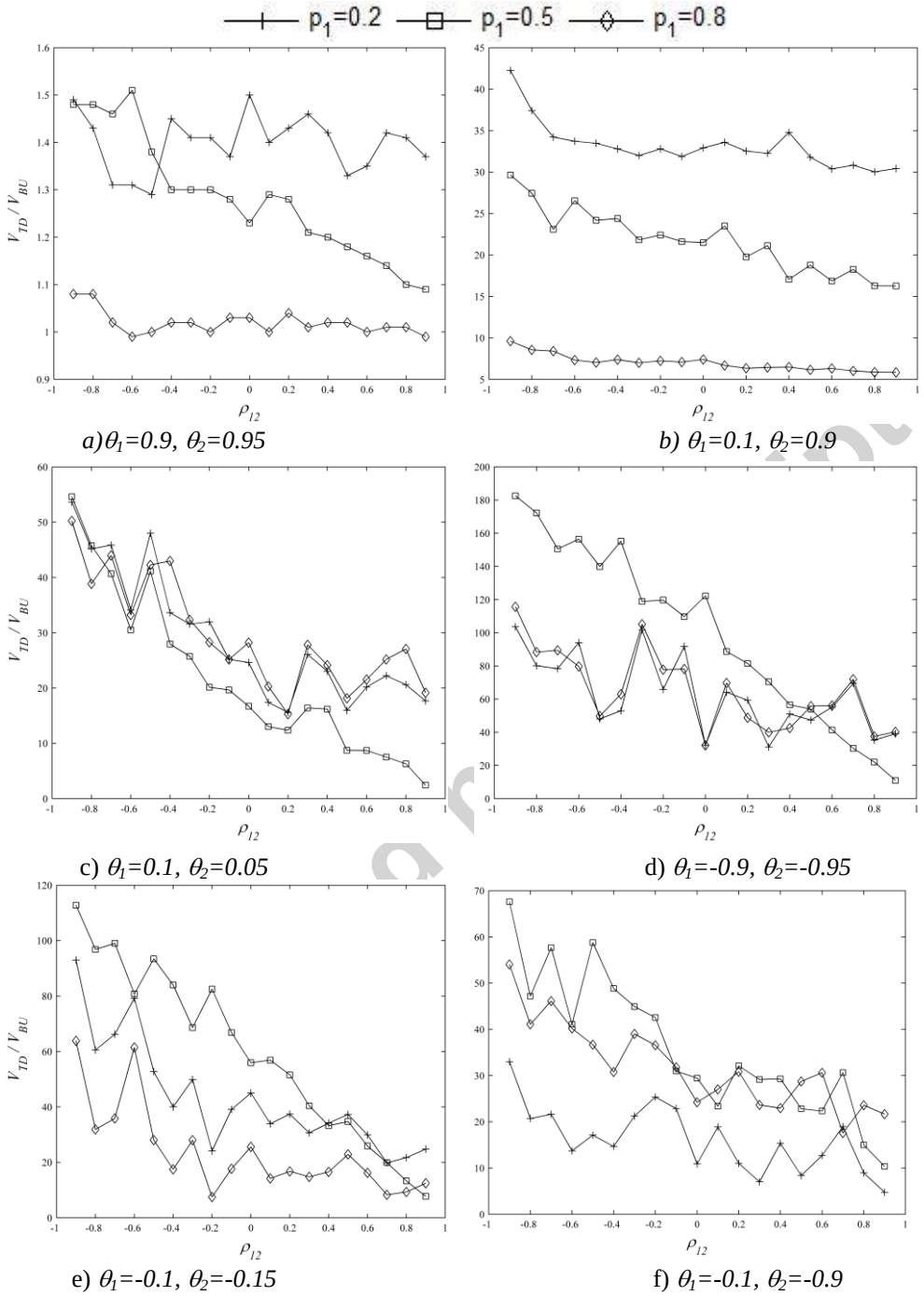
Our findings are somehow in agreement with some of the earlier studies in this area by Barnea and Lakonishok (1980) and Fliedner (1999) (although we do note that our results are not directly comparable to these studies as we analyse a non-stationary case). The analysis of Barnea and Lakonishok (1980) based on an empirical evaluation showed that positive cross-correlation contributes to the superiority of forecasts based on aggregate data (TD), which is also the case in our study. Fliedner (1999) used a simulation study to compare the performance of TD and BU in forecasting aggregate series where the two sub-aggregate items follow an MA(1) process. He found that TD dominated BU regardless of the values of the cross-correlation coefficient. They have not reported the values of θ_1 and θ_2 used in their study, so our interpretation is that this work considered only the opposing signs for θ_1 and θ_2. Should this be the case then these findings are in agreement with ours.

4.2. Comparison at the Sub-Aggregate Level

In this subsection we evaluate the relative performance of the TD and BU approaches in forecasting sub-aggregate demand when the moving average parameters are not necessarily identical. The simulation structure in terms of within and out-of-sample arrangements is as discussed in the previous sub-section. Under the BU approach, we generate 300 one step-ahead forecasts for each item individually using the optimal smoothing constant. Under the TD approach, we first sum up the demand of all sub-aggregate items to obtain aggregate series, we then produce the aggregate forecast and finally we break down (disaggregate) that to sub-aggregate forecasts by using proportional factors based on the historical contribution of each series. For each experiment, the ratio of the variance of forecast error is calculated as: $\sum_{i=1}^{2} Var\left(d_{i,t} - p_i * F_t\right) \Big/ \sum_{i=1}^{2} Var\left(d_{i,t} - f_{i,t}\right)$.

Figure 3 shows the ratio of the variance of forecast error of the TD over the BU approach at the sub-aggregate level for different values of θ_1, θ_2 , ρ_{12} and p_1 when the sub-aggregate items follow an ARIMA(0,1,1) process with non-identical moving average parameters ($\theta_1 \neq \theta_2$). Different levels of item proportion, p_i, are used to reflect the cases where the sub-

aggregate items contribute almost equally to the aggregate forecast and cases where one item dominates the aggregation process. The results show that the BU approach always outperforms TD in forecasting the sub-aggregate items regardless of the ρ_{12} and p_i.



a) θ_1*=0.9,* θ_2*=0.95*

b) θ_1*=0.1,* θ_2*=0.9*

c) θ_1*=0.1,* θ_2*=0.05*

d) θ_1*=-0.9,* θ_2*=-0.95*

e) θ_1*=-0.1,* θ_2*=-0.15*

f) θ_1*=-0.1,* θ_2*=-0.9*

Figure 3. Relative performance of TD and BU approaches in forecasting sub-aggregate items under different values of θ_1, θ_2, ρ_{12} and p_1.

In Figure 3 we show that by moving from a cross-correlation of -0.9 toward +0.9 the ratio of V_{TD}/V_{BU} is generally reduced but it always remains greater than 1 regardless of the cross-correlation coefficient and the items' contributory power.

When the cross-correlation and the moving average parameters, θ_1, θ_2, are highly positive, *i.e.* $\theta_1 \cong 0.99$, $\theta_2 \cong 0.99$ and $\rho_{12} \cong 0.99$, the ratio of V_{TD}/V_{BU} becomes close to one.

Figure 3a shows also that BU outperforms TD by a maximum of about 50% for highly negative cross-correlation; the rate of superiority of BU becomes very high when θ_1 and θ_2 are not highly positive (see Figure 3b, c, d). Widiarta et al. (2009) reported that when demand follows an MA(1) process, BU outperformed TD by a maximum of 4% when θ_1=0.3 and θ_2=-0.8 and the cross-correlation is negative. The superiority of BU at the sub-aggregate level can be attributed to the potentially high positive autocorrelation between demand periods. When the series follow an ARIMA(0,1,1) process, the autocorrelation is (highly) positive unless the moving average parameter takes high positive values, in which cases autocorrelation becomes close to zero. Generally, as θ_i moves from positive toward negative values, the autocorrelation between two consecutive observations $d_{i,t}$ positively increases, in addition to spanning higher time lags (not only a lag of 1). This makes it much more difficult to apportion the resulting aggregate forecast, *Ft*, to each item in the family based on the historical demand proportion, p_i. As a result, the performance of the TD approach is affected adversely. The performance of the BU approach, however, is not affected as it forecasts the demand for each item individually.

Our findings are in accordance with those previously reported in the academic literature. Widiarta et al. (2007) argued that when the sub-aggregate time series follow an AR(1) process and the value of the autocorrelation is high, there is a sharp deterioration in the relative performance of TD. Gordon et al. (1997) and Dangerfield and Morris (1992) used empirical data from the M-competition database and stated that BU dominated TD when forecasting the sub-aggregate time series. Weatherford et al.(2001) have shown that a purely disaggregate forecast (BU) strongly outperformed even the best aggregate forecast (TD) at the sub-aggregate level.

These results generally confirm our findings although we must note (as we did in the previous sub-section) that there is not a direct comparison between these studies and ours due to the consideration of a non-stationary ARIMA(0,1,1) time series process. Contrasting our results with those reported by Widiarta et al. (2007, 2009) on stationary MA(1) and AR(1) processes, we observe that the rates of superiority of the BU approach when the process is

non-stationary is much higher than the stationary case. When the demand follows a stationary AR(1) process, the maximum ratio of V_{TD}/V_{BU} equals to 2 and is obtained with series with high positive autocorrelation (high positive autoregressive parameter values ϕ_1, ϕ_2) and highly negative cross-correlation, while this ratio for the IMA(1,1) process is higher than 50.

5. EMPIRICAL ANALYSIS

In this section, we assess the empirical validity of our results on the comparative performance of the TD and BU approaches. In addition, the empirical performances of these approaches are compared to that of the optimal approach proposed by Hyndman et al. (2011), at both the aggregate and subaggregate levels. We first provide details of the empirical data available for the purposes of our investigation along with the experimental structure employed in our work. We then present the actual empirical results.

The demand dataset available for the purposes of our research consists of 103 weekly sales observations (i.e. it spans a period of two years) for 1,798 SKUs from a European grocery store. The *auto.arima* function of the forecast package in R has been used to identify the underlying ARIMA demand process for each series and to estimate the relevant parameters. This function uses a variation of the Hyndman and Khandakar (2008) algorithm which combines unit root tests, minimization of the Akaike's Information Criteria (*AICc*) and maximum-likelihood estimation (*MLE*) to identify an ARIMA(p,d,q) model. First, the number of differences *d* is determined using unit-root tests by applying repeated KPSS tests (Kwiatkowski et al., 1992). Then, the value of process orders, *p* and *q*, are chosen by minimizing the *AICc* after differencing the data *d* times. Please refer to Hyndman and Khandakar (2008) for a discussion on ARIMA identification methodology related issues.

Based on the identification process discussed above, it was found that around 24% of the series (424 series) may be represented by the process considered in this research, ARIMA(0,1,1). In Table 2 we summarize the characteristics of the SKUs relevant to our study by indicating the estimated parameters for the ARIMA(0,1,1) process. It is important to note that these results are sensitive to the modelling methodology being used to identify the series in the first place. If the methodology employed by the *auto-arima* function potentially identifies ARIMA(0,1,1) series incorrectly then our results will be obviously subject to relevant errors.

To facilitate a clear presentation, the estimated parameters are grouped in intervals and the corresponding number of SKUs is given for each such interval. The average θ value per

interval is also presented. This categorisation allows us to compare the empirical results with the theoretical findings. We must remark that the θ parameter values are all positive, except for two SKUs, and most of them take highly positive values. As such, the data do not cover the entire theoretically feasible range of the parameters.

Table 2. Processes present in the empirical data set

Group	θ intervals	Average of θ	No. of SKUs
1	[0.1,0.3[	0.2097	4
2	[0.3,0.4[	0.3652	8
3	[0.4,0.5[	0.4656	17
4	[0.5,0.6[	0.5591	32
5	[0.6,0.7[	0.6561	67
6	[0.7,0.8[	0.7503	108
7	[0.8,0.9[	0.8467	141
8	[0.9,1]	0.9534	47
Total number of SKUs:			424

The data series have been divided into two parts. The first part (within sample) consists of 70 time periods and is used in order to estimate the *SES* parameters. The second part consists of 33 time periods which are used to evaluate the performance of each approach (out-of-sample). The geometric mean (across SKUs) of the V_{TD}/V_{OP} and V_{BU}/V_{OP} ratios is considered for comparison purposes at the disaggregate level (where V_{OP} is the variance of the forecast errors resulting from the implementation of the optimal approach). Note that the ratio V_{TD}/V_{BU} can be directly deduced from the two variance ratios as $(V_{TD}/V_{OP}) / (V_{BU}/V_{OP})$.

The empirical results presented in Table 3 are shown for the same θ intervals considered in Table 2. With regards to the comparative performance of BU and TD, we can see that when the smoothing constant values are optimised for both approaches, the variance ratio is greater than 1 when the comparison is undertaken at the sub-aggregate level, whereas when the comparison is undertaken at the aggregate level, the difference between BU and TD is insignificant. This means that overall one can consider that the BU approach provides more accurate forecasts. Furthermore, when the smoothing constants used for BU and TD are equal, the ratio of V_{TD}/V_{BU} equals to 1 in the case of disaggregate demand forecasting. As discussed above, the moving average parameter θ is highly positive for most SKUs considered in this research. More than 85% of the SKUs have a moving average parameter greater than 0.6 (see Table 2). In addition, the sub-aggregate cross-correlation coefficients between SKUs vary

between -0.5 and +1; however most of those coefficients are positive. By referring to the detailed results of the simulation study we see that for this range of moving average parameter values, $0<\theta<1$, the BU approach performs better than TD at the subaggregate level. However, for the comparison at the aggregate level, the ratio of V_{TD}/V_{BU} is close to one, and the superiority of each approach depends on the cross-correlation. TD may outperform BU for highly positive cross-correlation. In an empirical context, the average of the variance of forecast error reduction may be as high as 1.01% when the comparison is performed at the aggregate level, while 70% variance error reduction may be achieved for the comparison at the sub-aggregate level.

Table 3. Empirical variance ratios for an ARIMA(0,1,1) process

Group	**θ intervals**	**Comparison Level**			
		Aggregate		***Disaggregate***	
		V_{TD}/V_{OP}	V_{BU}/V_{OP}	V_{TD}/V_{OP}	V_{BU}/V_{OP}
1	[0.1,0.3[	0.998	1.020	2.391	0.990
2	[0.3,0.4[	1.001	1.022	2.078	0.993
3	[0.4,0.5[	1.000	1.005	1.880	0.997
4	[0.5,0.6[	1.000	1.002	1.727	0.994
5	[0.6,0.7[	1.000	1.020	1.629	0.990
6	[0.7,0.8[	1.000	1.008	1.428	1.003
7	[0.8,0.9[	1.000	1.007	1.290	0.996
8	[0.9,1]	1.001	1.000	1.162	0.999
Average		**1.000**	**1.011**	**1.613**	**0.995**

With regards to the implementation of the optimal approach, the hierarchy structure of our data consists of two levels. At the top level we have aggregated all series to get one single series while at the bottom level we have 424 series, so we have 425 series in total for forecasting purposes. First, we generated forecasts for all 425 series using SES. Next, we have used the *combinef* function in the *hts* package (Hyndman et al., 2014) of R to reconcile these forecasts using the optimal method to obtain forecasts at both aggregate and sub-aggregate levels.

With regards to the comparative performance of the optimal approach (as reported based on its variance of forecast errors, V_{OP}), Table 3 shows that for comparison at the aggregate level, the optimal method performs better than BU; however there is no significant difference

between the performance of the optimal method and the TD approach. When the comparison is undertaken at the sub-aggregate level, the optimal method significantly outperforms TD, whereas BU works slightly better than the optimal method (the difference is less than 1%). Generally, for the empirical data used in this study, either the TD or the optimal method can be used for forecasting at the aggregate level, while at the sub-aggregate level, BU is preferable, although the optimal approach may also be used as the difference is less than 1%.

In Table 3, we report ratios of the variance of forecast errors in specific moving average parameter interval. In Table 4, we report collective performance across different possible (ranges of) moving average parameter values and we evaluate the impact of such values on the superiority of each approach. To do so we create a category containing groups 1, 2 and 3 that includes 29 SKUs; this is regarded as a category with the lowest values of θ. As we move from this category to groups 4, 5 and 6 the value of θ increases. We aggregate all these groups with group 8 that represents the highest values of θ. The ratios V_{BU}/V_{OP} and V_{TD}/V_{OP} are then presented for both levels of comparison.

Table 4. Empirical variance ratios by reporting performance across different groups (intervals of θ values)

Comparison Level		**Group**	**1,2,3**	**4**	**5**	**6**
Aggregate	V_{TD}/V_{OP}	**8**	1.001	1.001	1.000	1.000
	V_{BU}/V_{OP}		0.855	0.902	0.977	0.998
Subaggregate	V_{TD}/V_{OP}		1.484	1.435	1.399	1.366
	V_{BU}/V_{OP}		0.956	0.976	0.987	0.998

With regards to the comparative performance of BU and TD, the results indicate that when two groups with different moving average parameters are considered (Group 1,2,3 with 8) then the variance ratio V_{TD}/V_{BU} is high and as the θ values increases (tending towards the values covered by group 8) the ratio decreases. This implies that when we aggregate groups of SKUs with low and high θ values there is a greater benefit of using the BU approach in terms of accuracy. This is exactly what we have observed in the simulation results for 2 SKUs (one associated with a small and one with a high θ value). These empirical results generally confirm the findings of the simulation study. With regards to the performance of the optimal approach, Table 4 shows that for comparisons at the sub-aggregate level, the optimal method outperforms TD. However, BU performs better than the optimal approach. When

performances are contrasted at the aggregate level, both optimal and TD approaches perform equally well and BU outperforms them.

6. DISCUSSION, CONCLUSION AND FURTHER RESEARCH

In this paper we have evaluated analytically the effectiveness of the bottom-up (BU) and top-down (TD) approaches to forecasting aggregate and subaggregate demand when the subaggregate series follow a first order integrated moving average [ARIMA(0,1,1)] process. Forecasting was assumed to be relying upon a Single Exponential Smoothing (SES) procedure and the analytical results were complemented by a simulation experiment at both the aggregate and sub-aggregate level as well as experimentation with an empirical dataset from a European superstore.

Admittedly, the current fast changing market environments result in many demand processes being non-stationary in nature. Some empirical pieces of work discussed in subsection 2.1 confirm such a statement and provide support for the frequency with which ARIMA(0,1,1) processes are encountered in real world applications. In addition, SES is a most commonly employed forecasting procedure in industry and its application implies a non-stationary behaviour (SES is optimal for an ARIMA(0,1,1) process). Both BU and TD approaches are very useful in practice when dealing with Sales and Operations Planning systems in which forecasting is required at both aggregate and subaggregate levels. In summary, we feel that the problem setting we have considered is a very realistic one. Analytical and simulation developments were based on the consideration of the variance of forecast error for the TD and BU approaches and comparisons were undertaken at the aggregate level in the theoretical part of this work and at both the subaggregate and aggregate level in the simulation investigation. The conditions under which one approach outperforms the other were identified and the main findings can be summarized as follows:

- When the moving average parameter for all the subaggregate items is identical ($\theta_1=\theta_2=\ldots=\theta_N$), there is no significant difference between TD and BU in forecasting the aggregate level, as long as the optimal smoothing constant is used for both approaches. When the smoothing constant used for all the subaggregate items and the aggregate level is set to be identical ($\alpha=\alpha_{TD}$), TD and BU perform the same in forecasting the demand at the aggregate level regardless of the values of the moving average parameters and the cross-correlation between items. In addition when the

observations of the subaggregate items are highly auto-correlated (negative θ), the performance of BU and TD is also the same for all autocorrelation values.

- TD performs better than BU at the aggregate level when the subaggregate moving average parameters take relatively high values of an opposing sign and the cross-correlations between sub-aggregate items are (highly) positive. Otherwise, when cross-correlation is positive low or takes negative values, BU is preferable. Therefore, using the aggregate data to produce top level forecasts is preferable only if the sub-aggregate items follow similar series evolution with combination of high vs. low autocorrelation. The TD appears not to be very accurate when the sub-aggregate items consist of different patterns of fluctuation.
- BU outperforms TD when forecasting at the sub-aggregate level and when the smoothing constant is set to its optimal value for both approaches, regardless of the cross-correlation, the disaggregation weight and the values of the process parameters. It's not preferable to use the aggregate data to derive the individual forecasts, when the autocorrelation of subaggregate items is highly positive, in which cases subaggregate data provide more accurate forecasts. The degree of superiority of the BU approach for non-stationary processes is much higher compared to stationary ones.
- The performance of BU improves as the cross-correlation decreases, moving from positive toward negative values. For highly negative cross-correlation values BU is always preferable; this is generally true for comparisons at both the aggregate and sub-aggregate level.
- The benefit achieved by BU and TD for non-stationary demands in terms of forecast accuracy is higher than that associated with stationary cases.
- The optimal approach performs well at both levels of comparison as indicated by the empirical results. When considering ratios of the variance of forecast errors in particular moving average parameter intervals or different possible (ranges of) moving average parameter values, the optimal approach is superior as it performs as well as the BU at both levels but significantly better than the TD at the disaggregate level. It should be noted though that the optimal approach bears considerable relevance to many realistic cases when: i) more than two levels of hierarchy need to be considered, and ii) (more than one) various forecasting methods need to be employed.

Please also note that since the optimal approach is not always superior to the BU and TD approaches, the comparative performance between BU and TD needs to be carefully

considered in order to inform real world applications. In addition, and even when the optimal approach outperforms both BU and TD, the latter approaches still remain of explicit interest to practitioners (because of the simplicity characterising their implementation, their intuitive appeal and their support by most inventory software) but also to academics (because of the insights that the mathematical analysis of those cases may offer – in contrast with the optimal approach where the optimisation procedure hinders any explicit messages as to what is going on with the underlying properties of the series.) The major difficulty associated with the optimal approach is the computation of the reconciliation weights used to form a weighted average of forecasts at an individual node and the non-transparent nature of the regression analysis taking place. Another difficulty is the fact that forecasts from all levels need to be taken into account when producing the final (reconciled) forecasts which obviously increases the computational effort and managerial involvement beyond that required by either BU or TD.

If the practitioners require demand forecasts at the SKU level when demand is non-stationary (and highly autocorrelated) then it would always be preferable to use the BU approach. If a higher level demand forecast is needed then the BU approach should be considered when the individual items are associated with different patterns of evolution, and the TD or the optimal approach when they have the same patterns but are associated with different autocorrelation values. In addition, if one uses the same value of smoothing constants for both BU and TD, then both approaches perform the same in forecasting aggregate demand.

In this paper we have considered the case of non-stationary demand processes in conjunction with the SES forecast method to evaluate the comparative performance of TD and BU in forecasting aggregate and item level demand. Naturally, there are many other avenues for further research and the following possibilities should be very important in terms of advancing the current state of knowledge in this area.

- The consideration of more extensive datasets that cover the whole range of the process parameters should allow a better understanding of the comparative benefits of the TD and BU approach.
- The extension of the work described here to cover inventory/implication metrics would allow a linkage between forecasting and stock control. Cross-sectional aggregation is known to be very helpful in inventory applications and it is indeed being covered by relevant software packages. Further work into the interactions

between forecasting and stock control in a cross-sectional setting should add value to real world practices.

- The interface between (and the potential of combining) temporal and cross-sectional aggregation has received minimal attention both in academia and industry and is an issue that we are to take further in the next steps of our research. A *unified approach* – one that simultaneously considers choices of aggregation levels and frequency along multiple dimensions – would seem to be a valuable step in the right direction. The problem with the separation of the cross-sectional and temporal dimensions is that the *right level* of cross-sectional aggregation may vary across time frequencies and vice versa. Procedures that combine forecasts for a cross-sectional hierarchy, such as the optimal approach discussed in this paper, and procedures that combine forecasts over time frequencies, such as the multiple temporal-aggregation technique discussed by Kourentzes et al. (2014), may conceivably be *pooled* to form a holistic strategy for forecasting hierarchies (see also the Introduction of Tashman et al. (2015)).
- Finally, the analytical and empirical consideration of Integer ARMA (INARMA) processes offers a great opportunity for advancements in the area of aggregation. Such processes bear a considerable relevance to intermittent demands where the benefits of aggregation may be even higher due to the reduction of zero observations (Mohammadipour and Boylan, 2012).

REFERENCES

Acar, Y., Gardner, E.S., 2012. Forecasting method selection in a global supply chain. International Journal of Forecasting 28, 842-848.

Ali, M.M., Boylan, J.E., Syntetos, A.A., 2011. Forecast errors and inventory performance under forecast information sharing. International Journal of Forecasting.

Athanasopoulos, G., Ahmed, R.A., Hyndman, R.J., 2009. Hierarchical forecasts for Australian domestic tourism. International Journal of Forecasting 25, 146-166.

Barnea, A., Lakonishok, J., 1980. An Analysis of The Usefulness of Disaggregated Accounting Data For Forecasts of Corporate Performance. Decision Sciences 11, 17-26.

Bijvank, M., Vis, I.F.A., 2011. Lost-sales inventory theory: A review. European Journal of Operational Research 215, 1-13.

Box, G., Jenkins, G.M., Reinsel, G., 2008. Time Series Analysis: Forecasting & Control. John Wiley and sons.

Bozos, K., Nikolopoulos, K., 2011. Forecasting the value effect of seasoned equity offering announcements. European Journal of Operational Research 214, 418-427.

Chen, H., Boylan, J.E., 2007. Use of Individual and Group Seasonal Indices in Subaggregate Demand Forecasting. The Journal of the Operational Research Society 58, 1660-1671.

Chen , H., Boylan , J.E., 2008. Empirical evidence on individual, group and shrinkage seasonal indices. International Journal of Forecasting 24, 525-534.

Chien, C.-F., Chen, Y.-J., Peng, J.-T., 2008. Demand forecast of semiconductor products based on technology diffusion, Simulation Conference, 2008. WSC 2008. Winter, pp. 2313-2322.
Dangerfield, B.J., Morris, J.S., 1992. Top-down or bottom-up: Aggregate versus disaggregate extrapolations. International Journal of Forecasting 8, 233-241.
Erkip, N., Hausman, W.H., Nahmias, S., 1990. Optimal Centralized Ordering Policies in Multi-Echelon Inventory Systems with Correlated Demands. Management Science 36, 381-392.
Fliedner, G., 1999. An Investigation of Aggregate Variable Time Series Forecast Strategies with Specific Subaggregate Time Series Statistical Correlation. Computers and Operations Research 26, 1133-1149.
Gardner, E.S., Jr., 1990. Evaluating Forecast Performance in an Inventory Control System. Management Science 36, 490-499.
Gardner, E.S., Jr., 2006. Exponential smoothing: The state of the art-Part II. International Journal of Forecasting 22, 637-666.
Goh, C., Law, R., 2002. Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. Tourism Management 23, 499-510.
Gordon, T., Dangerfield, B., Morris, J., 1997. Top-down or bottom-up: Which is the best approach to forecasting? The Journal of Business Forecasting 16, 13-16.
Granger, C.W.J., Morris, M.J., 1976. Time series modeling and interpretation. Journal of the Royal Statistical Society. Series A, 246-257.
Graves, S.C., Willems, S.P., 2000. Optimizing Strategic Safety Stock Placement in Supply Chains. Manufacturing & Service Operations Management 2, 68-83.
Graves, S.C., Willems, S.P., 2008. Strategic Inventory Placement in Supply Chains: Nonstationary Demand. Manufacturing & Service Operations Management 10, 278-287.
Gross, C.W., Sohl, J.E., 1990. Disaggregation methods to expedite product line forecasting. Journal of Forecasting 9, 233-254.
Hyndman, R.J., Ahmed, R.A., Athanasopoulos, G., Shang, H.L., 2011. Optimal combination forecasts for hierarchical time series. Computational Statistics & Data Analysis 55, 2579-2589.
Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: The forecast package for R. Journal of Statistical Software, 26, 1-22.
Hyndman, R.J., Koehler, A.B., Ord, J.K., Snyder, R.D., 2008. Forecasting with Exponential Smoothing: The State Space Approach. Springer Verlag, Berlin.
Hyndman, R.J., Koehler, A.B., Snyder, R.D., Grose, S., 2002. A state space framework for automatic forecasting using exponential smoothing methods. International Journal of Forecasting 18, 439-454.
Hyndman, R.J., Wang, E., Lee, A., 2014. Package 'hts': Methods for analysing and forecasting hierarchical and grouped time series.
Kourentzes, N., Petropoulos, F., Trapero, J.R., 2014. Improving forecasting by estimating time series structural components across multiple frequencies. International Journal of Forecasting 30, 291-302.
Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics 54, 159-178.
Lapide, L., 2004. Sales and Operations Planning Part I: The Process. The Journal of Business Forecasting, 17-19.
Lapide, L., 2006. Top-Down & Bottom-Up Forecasting In S&OP. The Journal of Business Forecasting 25, 14-16.
Lee, H.L., P., P., Whang , S., 1997. Bullwhip effect in a supply chain. Sloan Management Review 38, 93-102.
Li, B., Wang, H.-W., Yang, J.-B., Guo, M., Qi, C., 2011. A belief-rule-based inventory control method under nonstationary and uncertain demand. Expert Systems with Applications 38, 14997-15008.
Lütkepohl, H., 1984. Forecasting contemporaneously aggregated vector ARMA processes. Journal of Business & Economic Statistics 2, 201-214.
Mahajan, S., Desai, O., 2011. Value of Information in a Serial Supply Chain under a Nonstationary Demand Process, Working Paper Series. Indian Institute of Management (IIMB), Bangalore
Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., Winkler, R., 1982. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. Journal of Forecasting 1, 111-153.

Martel, A., Diaby, M., Boctor, F., 1995. Multiple items procurement under stochastic nonstationary demands. European Journal of Operational Research 87, 74-92.
Mitchell, G., Niederhausen, M., 2010. On Replenishing Items with Seasonal Intermittent Demand. American Journal of Economics and Business Administration 2, 90-102.
Mohammadipour, M., Boylan, J., Syntetos, A., 2012. The Application of Product-Group Seasonal Indexes to Individual Products. Foresight: The International Journal of Applied Forecasting, 18-24.
Mohammadipour, M., Boylan, J.E., 2012. Forecast horizon aggregation in integer autoregressive moving average (INARMA) models. OMEGA 40, 703-712.
Orcutt, G.H., Watts, H.W., Edwards, J.B., 1968. Data aggregation and information loss. The American Economic Review 58.
Raghunathan, S., 2001. Information Sharing in a Supply Chain: A Note on its Value when Demand Is Nonstationary. Management Science 47, 605-610.
Sbrana, G., Silvestrini, A., 2013. Forecasting aggregate demand: Analytical comparison of top-down and bottom-up approaches in a multivariate exponential smoothing framework. International Journal of Production Economics.
Shang, K.H., 2012. Single-Stage Approximations for Optimal Policies in Serial Inventory Systems with Nonstationary Demand. Manufacturing & Service Operations Management.
Shlifer, E., Wolff, R.W., 1979. Aggregation and Proration in Forecasting. Management Science (pre-1986) 25, 594.
Syntetos, A.A., 2001. Forecasting of intermittent demand, Ph D, Business School, Buckinghamshire Chilterns University College, Brunel University, London, UK.
Tashman, L., 2015. Guide on 'Techniques for forecasting product and temporal hierarchies. FORESIGHT: The international journal of applied forecasting.
Taylor, J.W., 2003. Exponential smoothing with a damped multiplicative trend. International Journal of Forecasting 19, 715-725.
Treharne, J.T., Sox, C.R., 2002. Adaptive Inventory Control for Nonstationary Demand and Partial Information. Management Science 48, 607-624.
Tunc, H., Kilic, O.A., Tarim, S.A., Eksioglu, B., 2011. The cost of using stationary inventory policies when demand is non-stationary. Omega 39, 410-415.
Weatherford, L.R., Kimes, S.E., Scott, D.A., 2001. Forecasting for hotel revenue management: Testing aggregation against disaggregation. The Cornell Hotel and Restaurant Administration Quarterly 42, 53-64.
Widiarta, H., Viswanathan, S., Piplani, R., 2007. On the effectiveness of top-down strategy for forecasting autoregressive demands. Naval Research Logistics (NRL) 54, 176-188.
Widiarta, H., Viswanathan, S., Piplani, R., 2008. Forecasting item-level demands: an analytical evaluation of top–down versus bottom–up forecasting in a production-planning framework. IMA Journal of Management Mathematics 19, 207-218.
Widiarta, H., Viswanathan, S., Piplani, R., 2009. Forecasting aggregate demand: An analytical evaluation of top-down versus bottom-up forecasting in a production planning framework. International Journal of Production Economics 118, 87-94.
Zotteri, G., Kalchschmidt, M., Caniato, F., 2005. The impact of aggregation level on forecasting performance. International Journal of Production Economics 93-94, 479-491.