

Output-error state-space identification of vibrating structures using evolution strategies: a benchmark study

Vasilis K. Dertimanis*

*Department of Civil Engineering and Geomatics, Faculty of Engineering and Technology,
Cyprus University of Technology, Limassol 3603, Cyprus, P.O. Box 50329*

(Received September 11, 2013, Revised April 30, 2014, Accepted June 25, 2014)

Abstract. In this study, four widely accepted and used variants of Evolution Strategies (ES) are adapted and applied to the output-error state-space identification problem. The selection of ES is justified by prior strong indication of superior performance to similar problems, over alternatives like Genetic Algorithms (GA) or Evolutionary Programming (EP). The ES variants that are being tested are (i) the (1+1)-ES, (ii) the $(\mu/\rho+\lambda)$ - σ -SA-ES, (iii) the (μ_t, λ) - σ -SA-ES, and (iv) the (μ_w, λ) -CMA-ES. The study is based on a six-degree-of-freedom (DOF) structural model of a shear building that is characterized by light damping (up to 5%). The envisaged analysis is taking place through Monte Carlo experiments under two different excitation types (stationary / non-stationary) and the applied ES are assessed in terms of (i) accurate modal parameters extraction, (ii) statistical consistency, (iii) performance under noise-corrupted data, and (iv) performance under non-stationary data. The results of this suggest that ES are indeed competitive alternatives in the non-linear state-space estimation problem and deserve further attention.

Keywords: structural identification; evolution strategy; optimization; state-space

1. Introduction

As the design and construction of modern structures is constantly evolved to cover more complex and interdisciplinary functions, the evaluation of the actual structural performance is vital for a variety of reasons that include malfunction detection, health monitoring (Casciati 2010) and life-cycle management and assessment (Faravelli and Casciati 2004). In this respect, the importance of data-driven structural modeling in modern engineering has already been appreciated.

Identification of parametric models can take place in terms of both external (transfer function) and internal (state-space) descriptions of a structural system. The former has been traditionally the kernel of the envisaged research and many methods have been proposed for the estimation of discrete-time transfer functions, including linear least squares (Verhaegen and Verdult 2007, Huang 2001), linear multi-stage methods (Fassois 2001), maximum likelihood and nonlinear least squares (Ljung 1999).

*Corresponding author, Researcher, E-mail: vasileios.dertimanis@cut.ac.cy

The investigation on state-space parametric identification is relatively more recent and the main outcome of the research is the introduction of subspace methods for the estimation of system matrices (Verhaegen and Dewilde 1992, Van Overschee and De Moore 1996, Juang and Pappa 1985). While their significance has already been assessed in the relevant literature (Katayama 2006), subspace methods are suboptimal and, while they are considerably simple and fast, they are subject to potential inaccuracies, especially at low signal-to-noise-ratios (Viberg 1995).

Maximum likelihood estimation, on the other hand, leads to optimum solutions but suffers from problems related to nonlinear optimization, such as instabilities, local extrema and heavy dependence to initial “guesses” (Fassois 2001). In the state-space model case these problems are even more severe due to the inherent over-parameterizations that lead to surjective mappings (Mc Kelvey and Helmersson 1997). To tackle this problem, researchers have developed gradient projection methods (Mc Kelvey *et al.* 2004, Bergboer *et al.* 2002), with good results. However, the proposed algorithms are still amenable to initial “guesses” that affect convergence, while assessments in structural systems are still missing.

An alternative path towards the nonlinear estimation problem is the application of stochastic optimization algorithms (Fleming and Purshouse 2002). Various variants of the latter have already been applied to the estimation of transfer functions for structural identification problems (Dertimanis *et al.* 2003, Koulocheris *et al.* 2008), since the seminal paper by Kristinsson and Dumont (1992). However, these studies are limited only to SISO/MIMO transfer functions and corresponding investigations on state-space models have not been reported thus far.

Although their application in conventional, time-domain, parametric identification problem is somehow limited, stochastic optimization algorithms have been applied in inverse problems, including system identification (Franco *et al.* 2004) and structural health monitoring (Casciati 2008), which do not impose a certain model parameterization in transfer function or state-space format. Such approaches usually formulate a specific optimization problem in terms of physical parameters, (mass, stiffness and damping), which is then directly tackled by the adopted optimization algorithm through the establishment of a corresponding objective (fitness) function (Casciati 2008). A number of studies have been reported in this field that employ the use, among others, of GA (Zhang *et al.* 2010a, Zhang *et al.* 2010b), ES (Franco *et al.* 2004) and Differential Evolution (Tang *et al.* 2008). However, these approaches generally presume availability of the internal structure of the second-order structural equation, which indirectly implies knowledge of DOF, information that in many cases is not available.

Under this setting, the purpose of this study is to investigate whether stochastic optimization algorithms constitute an effective alternative to the state-space estimation problem. Following previous studies (Koulocheris *et al.* 2003), at which strong indication about the superior performance of ES against GA and EP has been observed for this class of optimization problems, four ES variants are selected and applied. The variants are:

- The (1+1)-ES: the simplest ES variant with just one parent and one offspring.
- The $(\mu/\rho+\lambda)$ - σ -SA-ES: the multi-membered ES with μ parents and λ offspring, in which recombination is performed using $1 \leq \rho \leq \mu$ parents. Mutation operates on both the individuals and their associated variances, in a self-adaptive manner that depends on certain externally selected learning parameters. The population for the next iteration is formulated by selecting the best μ individuals out of the $\mu+\lambda$ (plus scheme).
- The (μ_l, λ) - σ -SA-ES: similar variant like the previous one, except that the recombination operator is global intermediate ($\rho=\mu_l$) and that the population for the next iteration is formulated by selecting the best μ individuals out of the λ only (comma scheme).

- The (μ_w, λ) -CMA-ES: a second generation ES, in which each offspring is a mutated version of one centroid (produced by weighted global intermediate recombination) and where mutations are drawn from a multivariate Gaussian distribution.

The above algorithms are compared—through Monte Carlo experiments—over (i) accuracy of modal estimates, (ii) statistical consistency, (iii) performance under noise-corrupted vibration data and (iv) performance under non-stationary vibration data.

The benchmark study is based on a linear model of a six-story shear building subject to uniaxial horizontal earthquake excitation (ground acceleration). Despite its phenomenal simplicity, this model constitutes a very good test case, since it contains structural frequencies over a relatively short band, it is characterized by light damping (up to 5% for the first and the last mode, see Section 2) and its state-space description results in a parameterization vector of size $p = 234$, which renders the corresponding optimization problem quite demanding.

The rest of the paper is organized as follows: the structural identification problem is described in Section 2 and the four ES variants are outlined in Section 3. Section 4 contains the Monte Carlo benchmark analysis results and, finally, in Section 5 the results are summarized and discussed and some final remarks are given.

2. The structural identification problem

2.1 The simulation structure

Fig. 1 displays a simplified six-story shear building model subject to uniaxial earthquake excitation. Each story is modeled as a lumped mass, while the vertical columns are modeled as massless springs with equivalent stiffness. The equation of motion for the structure of Fig. 1 is

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{K}\mathbf{u}(t) = -\mathbf{M}\mathbf{r}\ddot{x}_g(t) \quad (1)$$

where \mathbf{M} and \mathbf{K} are the mass (diagonal) and stiffness (symmetric and positive definite) matrices, $\mathbf{u}(t) = [u_1(t) \ u_2(t) \ u_3(t) \ \dots \ u_6(t)]^T$ is the vector of relative (to the foundation) displacements of every story, \mathbf{r} is a vector of equal size to that of \mathbf{u} filled with ones, and $\ddot{x}_g(t)$ is the earthquake acceleration. To take into account energy dissipation, a corresponding velocity term has been added to Eq. (1), with the damping matrix obeying the Rayleigh form (α and β are real constants)

$$\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K} \quad (2)$$

Fig. 2 and Table 1 illustrate the mode shapes and the vibration modes, respectively, for the shear building of Fig. 1. The natural frequencies are contained within the $[1.5 \ 13.5]$ Hz band and the Rayleigh parameters have been chosen such, that the first and the last mode of vibration exhibit 5% damping ratio.

2.2 The identification problem

Given a finite number of earthquake excitation / structural response samples, the identification problem pertains to the estimation of the structural vibration modes. In a model-based framework,

this problem can be reduced to the estimation of a discrete-time state-space model of the form

$$\xi[t+1] = \mathbf{A}_d \xi[t] + \mathbf{B}_d \ddot{x}_g[t] \quad (3a)$$

$$\mathbf{y}[t] = \mathbf{G}_d \xi[t] + \mathbf{D}_d \ddot{x}_g[t] \quad (3b)$$

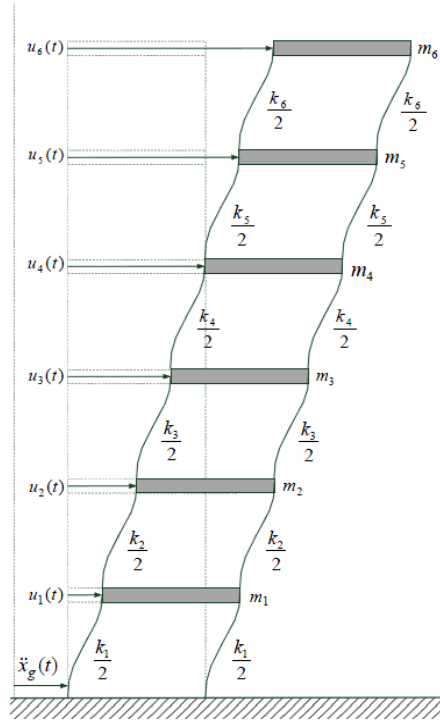


Fig. 1 Sketch of a six-story shear building model subject to single-axis horizontal earthquake excitation

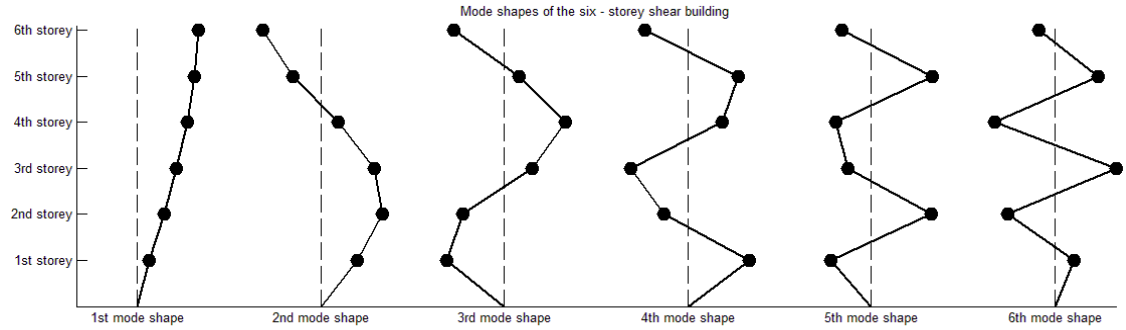


Fig. 2 Mode shapes of the six-story shear building model

Table 1 Physical and modal parameters of the building

Physical Space			Modal Space		
Story	m_j (Mgr)	k_j (kN/m)	Mode	f_n (Hz)	ζ_n (%)
1	100	200000	1	1.72	5.00
2	80	150000	2	5.00	3.19
3	80	150000	3	7.84	3.58
4	80	150000	4	10.17	4.13
5	80	150000	5	12.06	4.64
6	80	150000	6	13.33	5.00

the output of which closely resembles the measured vibration response. In Eq. (3(a)) $\xi[t]$ corresponds to the state vector and \mathbf{A}_d and \mathbf{B}_d to the state and input matrices, respectively. In Eq. (3(b)) $\mathbf{y}[t]$ refers to the structural vibration output and \mathbf{G}_d , \mathbf{D}_d to the output and feedforward matrices, respectively. The objective function of the induced optimization problem is given by

$$f(\mathbf{p}) = \mathbf{E}^T \mathbf{E} \quad (4)$$

where

$$\mathbf{E} = [\mathbf{e}^T[1] \mathbf{e}^T[2] \cdots \mathbf{e}^T[N]]^T \quad (5)$$

N is the length of the available data samples and $\mathbf{e}[t] = \mathbf{y}_m[t] - \mathbf{y}[t]$ is the prediction error between the measured and the predicted structural response at time t .

Having such a model adequately estimated, the identification of the vibration modes corresponds to the solution of the eigenvalue problem for the state matrix \mathbf{A}_d . It is noted that the discrete-time state-space model of Eq. (3) can be viewed as the digital equivalent of a continuous-time representation, as the latter results from Eq. (1) (Landau and Zito 2006).

Within the current study, the identification problem just described is handled under some underlying assumptions:

- (A1) The full parameterization is employed: all entries of the state-space matrices are entered into the parameter vector \mathbf{p} .
- (A2) All uncertainties due to noises acting on the system are lumped together as an additive disturbance at the output vector (Verhaegen and Verdult 2007). This means that the applied estimation procedure belongs to the class of output-error methods;
- (A3) The order of the state-space model is known a-priori: since the main purpose is to obtain a first insight about the behavior of the ES variants, no order estimation procedure is utilized;
- (A4) Measurements from all degrees-of-freedom are available.

As the above assumptions imply, the main focus of the current study is an initial investigation about the applicability and the performance of the selected ES variants, without imposing additional problem specific difficulties, as, for example, unknown state order, limited sensor availability, colored noise, etc. The latter are indeed very critical issues that pending to be

examined in a later study.

3. Summary of the applied ES variants

The family of ES consists one of the three most accepted variations of evolutionary algorithms (with GA and EP being the other two), which attempt to improve the quality of a parameter vector by using a series of operators that simulate natural evolution. In ES these operators are the recombination, mutation and selection ones and they act on a set of parent and offspring individuals in a way that is dictated by a specific ES variant. In general, the performance of ES is controlled by a set of endogenous parameters called the strategy parameters, which are also amenable to adaptation during the evolution cycle. For an excellent introduction to the fundamental notions of ES the reader is referred to Beyer and Schwefel (2001). In what follows, a brief introduction to the selected ES variants is taking place.

3.1 The (1+1)–ES

This is the simplest ES variant, which consists of just one parent and one offspring. The (1+1)–ES initializes with a random solution that is evaluated before the offspring is created. At every iteration, the offspring is created from the parent by mutating the latter using the multivariate normal distribution. That is, if \mathbf{x}_p and \mathbf{x}_o are the parent and offspring vectors, respectively, it follows that

$$\mathbf{x}_o = \mathbf{x}_p + \sigma \mathbf{z}, \quad \mathbf{z} \in N(\mathbf{0}, \mathbf{I}) \quad (6)$$

where σ is the stepsize parameter. The generated offspring is subsequently evaluated and replaces the parent if it produces a better fitness (in the current study, lower value of the objective function). Obviously, the most important parameter during the evolution process is the stepsize. It is updated during every $\max(5, p)$ iterations by implementing the 1/5 success rule (Beyer and Schwefel 2001): if the success rate is higher than 1/5th, the stepsize is increased; if the success rate is lower than 1/5th it is decreased; else, if the success rate is equal to 1/5th it remains unaltered.

Despite its simplicity and its reported poor performance on ill-conditioned and multimodal problems (Auger 2009), the (1+1)–ES is still, more than 40 years after its introduction, very useful. In fact, in widely used benchmark problems of stochastic optimization experiments, like the simple sphere, its serial efficiency is very hard to compete (Arnold 2006).

3.2 The family of $(\mu/\rho +, \lambda)$ – σ –SA–ES

A natural extension of the (1+1)–ES is the introduction of multi-membered versions, where the parents and the offspring are more than one. Thus, in the family of $(\mu +, \lambda)$ –ES (also referred to as canonical ES versions), λ offspring are generated from μ parents ($\mu < \lambda$), on the basis of recombination and mutation operators, during every generation. The selection operator determines the parents of the next generation by employing a completely deterministic process that chooses the best μ individuals out of the λ (comma scheme), or the best μ individuals out of the $\mu+\lambda$ (plus scheme).

Following the standard notation and terminology (Beyer and Schwefel 2001), an individual in the $(\mu +, \lambda)$ –ES is represented as

$$\mathbf{a}_k := \{\mathbf{x}_k, \mathbf{s}_k, f(\mathbf{x}_k)\} \quad (7)$$

where \mathbf{x}_k is the object parameter vector to be optimized, \mathbf{s}_k a vector of strategy parameters (a generalization of the stepsize in the (1+1)-ES) and $f(\mathbf{x}_k) : \mathfrak{R}^p \rightarrow \mathfrak{R}$ the objective function value. ES initializes with a parent population of individuals, $B_0 = \{\mathbf{a}_1, \mathbf{a}_1, \dots, \mathbf{a}_\mu\}$, the object parameter vectors of which are taken from the uniform distribution. The parental individuals are recombined to produce λ offspring, by applying the corresponding operator. In general, the recombination operator acts to a parent family of ρ individuals, which are randomly chosen from the tank of μ parents, with $1 \leq \rho \leq \mu$. The type of recombination that takes place in this study is the intermediate one, which calculates the centroid vector (mean) from the ρ parents. Notice that the recombination refers to the strategy parameters as well.

The recombined individuals are subsequently mutated in a similar way to that of the (1+1)-ES. Just as in the recombination operator, both the object parameters and the strategy parameters are included in the mutation process. In fact, in order to achieve co-evolution of object and strategy parameters, the latter are mutated first, so the mutation operator is described by the following set of equations

$$\forall k = 1 : \lambda, \quad \mathbf{a}_k \leftarrow \left\{ \begin{array}{l} \mathbf{s}_k = [\sigma_1 \dots \sigma_p] = e^{\tau_0 N(0,1)} [\sigma_1 e^{\tau N_1(0,1)} \dots \sigma_p e^{\tau N_p(0,1)}] \\ \mathbf{x}_k = \mathbf{x}_p + \mathbf{s}_k N(\mathbf{0}, \mathbf{I}) \\ f_k = f(\mathbf{x}_k) \end{array} \right\} \quad (8)$$

The scheme described by Eq. (6) is referred to as σ -self adaptation (SA) and it is functional without imposing external controls on the strategy parameters; self-adaptation depends on the, so called, learning parameters τ_0 and τ , which adopt certain values according to theoretical and experimental investigations (Beyer and Schwefel 2001).

The σ -SA-ES are very powerful stochastic optimization algorithms, according to performance assessment results (Arnold and Beyer 2002) and they have been used for parametric identification problems (Koulocheris *et al.* 2003, Koulocheris *et al.* 2004). In the current study, two variants are implemented and assessed: the (15/2+100)- σ -SA-ES and the (10_i,100)- σ -SA-ES (with the subscript denoting global intermediate recombination).

3.3 The (μ_w, λ) -CMA-ES

The Covariance Matrix Adaptation (CMA) ES variant (Hansen and Ostermeier 2001) belongs to the second generation ES. Essentially, it is an (μ_w, λ) -ES that uses weighted global intermediate recombination: each offspring is generated by the same parent, which is computed as the weighted centroid of μ individuals, that is

$$\langle \mathbf{x} \rangle_w = \sum_{j=1}^{\mu} w_j \mathbf{x}_j, \quad \sum_j w_j = 1 \quad (9)$$

The Eq. (6) of the original σ -SA-ES is substituted by

$$\forall k = 1 : \lambda, \quad \mathbf{a}_k \leftarrow \begin{cases} \mathbf{z}_k \leftarrow \mathbf{C}^{1/2} N(\mathbf{0}, \mathbf{I}) \\ \mathbf{x}_k = \langle \mathbf{x} \rangle_w + \sigma \mathbf{z}_k \\ f_k = f(\mathbf{x}_k) \end{cases} \quad (10)$$

and shows that each offspring is a mutated version of the centroid and that mutations are drawn from the $\sigma^2 N(\mathbf{0}, \mathbf{C})$ multivariate Gaussian distribution. This mechanism has the effect of adapting the mutation distribution to the local topology of the landscape (Willem 2012), causing the mutations to be taken along gradient directions and leading to significant improvement of the convergence speed. To successfully achieve this, both the covariance matrix \mathbf{C} and the stepsize σ must be properly adapted. This is taking place by proper CMA operations for the former quantity and cumulative stepsize adaptations (CSA) for the latter (Hansen and Ostermeier 2001).

The CMA-ES lies in the state of the art of evolutionary computation and experimentation has shown that it can reliably adapt topologies and improve convergence rates (Hansen 2006, Hansen and Ostermeier 2007). However, it is dependent to a (quite large) number of exogenous parameters, for which little theoretical knowledge is available and exhibits problems when dealing with populations of large size (Beyer and Sendhoff 2008). Here, the (10_w,20)-CMA version is utilized.

4. The structural identification experiments

4.1 Description

The ES variants discussed in Section 3 are now implemented to the structural identification problem described in Section 2.2., through Monte Carlo analysis that consists of 100 independent estimation experiments for each issue examined. During every such iteration of the applied Monte Carlo analysis, all the stochastic algorithms are initialized at the same random state, in order to prevent comparison under uncontrolled random effects. The selected exogenous parameters for each ES variant are depicted in Table 2, at which \mathbf{p}_0 denotes a subspace vector estimate (see below). For the CMA-ES, all underlying exogenous parameters are selected as described in (Hansen and Ostermeier 2001).

The simulation data have been obtained through the discretization of the structural equation (in fact its state-space representation) into the model of Eq. (3), at a sampling period $T_s = 0.01s$, using the zero-order hold (Landau and Zito 2006). The parameter vector \mathbf{p} of this discrete state-space model serves as the basis for all subsequent comparisons (where applicable).

Two estimation cases are considered, in respect to the type of the applied excitation signal: in the former, a zero-mean, stationary, Gaussian white noise process with standard deviation $\sigma_{xg} = 0.3m/s^2$ is applied as earthquake acceleration. In the latter the Northridge earthquake is utilized.

In order to make a comparison between the set of estimated state-space models and the original one, they are both transformed into the output normal form (Verhaegen and Verdult 2007), which provides a unique minimal parameterization. Then for each ES variant the following quantity is estimated

$$MSE = \frac{1}{100} \sum_{j=1}^{100} (\boldsymbol{\theta}_j - \boldsymbol{\theta})^T (\boldsymbol{\theta}_j - \boldsymbol{\theta}) \quad (9)$$

which corresponds to the mean square error of the parameters. In Eq. (9), $\boldsymbol{\theta}$ is the vector of the output normal form parameterization of the true state-space model and $\boldsymbol{\theta}_j$ the one of the estimated state-space model during Monte Carlo experiment j . Care must be taken to avoid confusion between the full parameterization and the output normal form parameterization: the former refers to the parameter vector to be optimized (denoted as \mathbf{p}_j), while the latter to the one of the output normal form parameterization, as it results from an *already estimated* state-space model.

It must be noted that primitive investigations about the applicability of the ES variants to the state-space estimation problem consisted of default values (Hansen and Ostermeier 2001, Beyer and Schwefel 2001) for the exogenous parameters, including the stepsizes and the initial intervals of the parental population. However, the results were extremely poor, in the sense that all methods failed to converge at a feasible solution point. That is why an initial parameter vector was estimated, by utilizing the deterministic subspace method described in paragraph 9.2.4 of Verhaegen and Verdult (2007). This vector is then used as a centroid individual, on the basis of which all mutations were employed.

Table 2 Selected exogenous parameters for each ES variant. For the actual values refer to the corresponding Section

Parameter	(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)–CMA
Initial population interval	$\mathbf{p}_0 \pm \sigma_0$	$\mathbf{p}_0 \pm \sigma_0$	$\mathbf{p}_0 \pm \sigma_0$	\mathbf{p}_0
Initial stepsize	n/a	σ_0	σ_0	σ_0
Stepsize low bound	σ_t	σ_t	σ_t	σ_t
Objective function threshold	f_t	f_t	f_t	f_t
Max. Function Evaluations	10^5	10^5	10^5	10^5

4.2 Case I: stationary excitation

4.2.1 Noise-free estimation

In the noise-free case, 1000 samples (per channel) of excitation–response data are utilized to the estimation of state-space models of Eq. (3), using the ES variants with the parameters of Table 2, where $\sigma_0 = 10^{-6}$ and $\sigma_t = f_t = 10^{-8}$. The results of the Monte Carlo analysis are expanded over Figs. 3 and 4 and Tables 3–6. Overall, the algorithms exhibit very good performance in all the examined issues, with the (10_I, 100)–σ–SA–ES showing the best results. However, the very satisfying statistical consistency is apparent in the estimates of the vibration modes (Table 6), which result identical to their theoretical counterparts at an almost negligible standard deviation. ES variant specific remarks are:

(1+1)–ES: as illustrated in Fig. 3, where the evolution of the (log₁₀) objective function vs.

evaluation times is plotted for all the 100 individual experiments, the algorithm is almost stuck for the first 50000 function evaluations and then starts to descend, reaching asymptotically the -5.5 line. In 95% of the experiments the algorithm terminates due to stepsize convergence and in the rest 5% due to objective function evaluation threshold, achieving a mean improvement of approximately 29% (Table 4), in comparison to the subspace estimate. The (1+1)-ES shows remarkable statistical consistency and this is verified from all the indicated parameters: the histogram of Fig.4, the standard deviation of the improvement (Table 4), the trace of the MSE estimate (Table 5), as well as the extraction of the vibration modes (Table 6).

(15/2+100)- σ -SA-ES: this ES variant exhibits considerably better performance than the previous one, managing to return better estimates in all the individual experiments, as a result of constantly decreasing iterations (see Fig. 3). The mean value of the best objective function lies in the area of -6.5 (observe the corresponding histogram of Fig. 4), yet, at a considerably higher deviation. In all experiments the algorithm terminates due to objective function evaluation threshold and achieves a mean improvement of more than 48% (Table 4), in comparison to the subspace estimate. The statistical consistency of this ES variant is not as remarkable as before, the vibration modes (Table 6) are, however, as accurate as the ones of the (1+1)-ES.

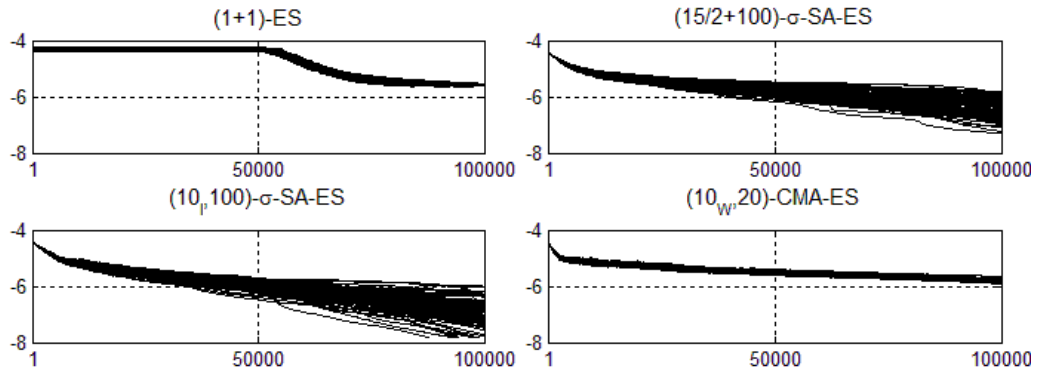


Fig. 3 $\log_{10}(f)$ values of each ES variant vs. objective function evaluations (stationary excitation case, noise-free)

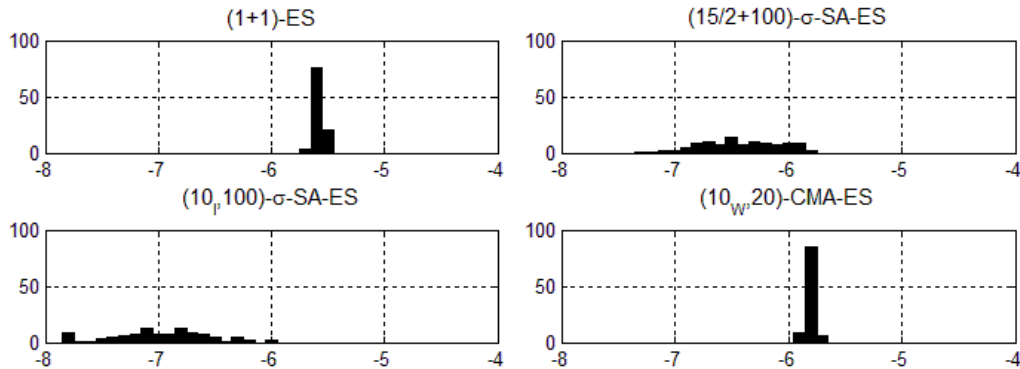


Fig. 4 Histograms of the final $\log_{10}(f)$ values of each ES variant at the end of every Monte Carlo iteration (stationary excitation case, noise-free)

Table 3 Percentage termination reason (stationary excitation case, noise-free)

Termination reason	(1+1)	(15/2+100)	(10 _l , 100)	(10 _l , 20)–CMA
Stepsize low bound	95%	0%	0%	0%
Objective function threshold	0%	0%	7%	0%
Max. Function Evaluations	5%	100%	93%	100%

Table 4 Percentage improvement from the subspace estimate (mean value \pm standard deviation, stationary excitation case, noise-free)

(1+1)	(15/2+100)	(10 _l , 100)	(10 _l , 20)–CMA
28.92% \pm 0.87%	48.34% \pm 7.96%	60.83% \pm 9.88%	33.86% \pm 0.85%

Table 5 Trace of the MSE of the output normal form parameterization (Eq. (9), stationary excitation case, noise-free)

(1+1)	(15/2+100)	(10 _l , 100)	(10 _l , 20)–CMA
1.45×10^{-6}	8.72×10^{-6}	1.22×10^{-5}	1.34×10^{-6}

Table 6 Estimated structural vibration modes (mean value, stationary excitation case, noise-free). Standard deviations were estimated in the order of 10-4 and therefore they're not shown

Mode	(1+1)		(15/2+100)		(10 _l ,100)		(10 _w ,20)–CMA	
	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)
1	1.72	5.00	1.72	5.00	1.72	5.00	1.72	5.00
2	5.00	3.19	5.00	3.19	5.00	3.19	5.00	3.19
3	7.84	3.58	7.84	3.58	7.84	3.58	7.84	3.58
4	10.17	4.13	10.17	4.13	10.17	4.13	10.17	4.13
5	12.06	4.64	12.06	4.64	12.06	4.64	12.06	4.64
6	13.33	5.00	13.33	5.00	13.33	5.00	13.33	5.00

(10_l, 100)- σ -SA-ES: the comma version of the self-adaptive ES returns the best results in terms of mean objective function values and percentage improvement to the subspace estimate. From Figs. 3 and 4 it is apparent that the mean value of the best objective function lies in the area of -7 (which corresponds to an order of magnitude smaller than the other variants, see Fig. 4), with higher deviation than before (about 10%). This is the only case where 7% of the experiments are terminated due to objective function threshold (10^{-8}) and the mean improvement over the subspace estimate climbs up to 61% (Table 4). Statistical consistency is good, but worse than the plus version. The vibration modes (Table 6) are, again, as accurate as the ones of the previous two variants.

(10_w, 20)-CMA-ES: in view of both the statistical consistency and the quality of the final estimate, the CMA-ES returns the best results. Here the consistency is as remarkable as the one of the first ES variant and the mean value of the best objective function approximates -6. In all experiments the algorithm terminates due to objective function evaluation threshold and the mean improvement in comparison to the subspace estimate is about 34% (Table 4). The CMA-ES returned the minimum MSE estimate (Table 5) and very accurate vibration modes (Table 6) as well.

4.2.1 Noise-corrupted estimation

In the stationary noise-corrupted case, white noise is added to the structural responses of each story at 5% noise-to-signal ratio (standard deviations). Again, 1000 samples (per channel) of excitation-response data are utilized to the estimation of state-space models of Eq. (3), using the ES variants with the parameters of Table 2, where now $\sigma_0 = 10^{-4}$ and $\sigma_t = f_t = 10^{-8}$.

The results of the Monte Carlo analysis are depicted in Figs. 5 and 6 and Tables 7-10. As expected, all ES variants do not manage to reproduce the objective function values reached in the noise-free case, while in all cases the algorithms are terminated due to the maximum objective function evaluation threshold. The mean percentage improvement from the subspace estimate is significantly lower (Table 4), starting from 7.39% for the (15/2+100)- σ -SA-ES and reaching up to 9.06%, again for the (10_l, 100)- σ -SA-ES.

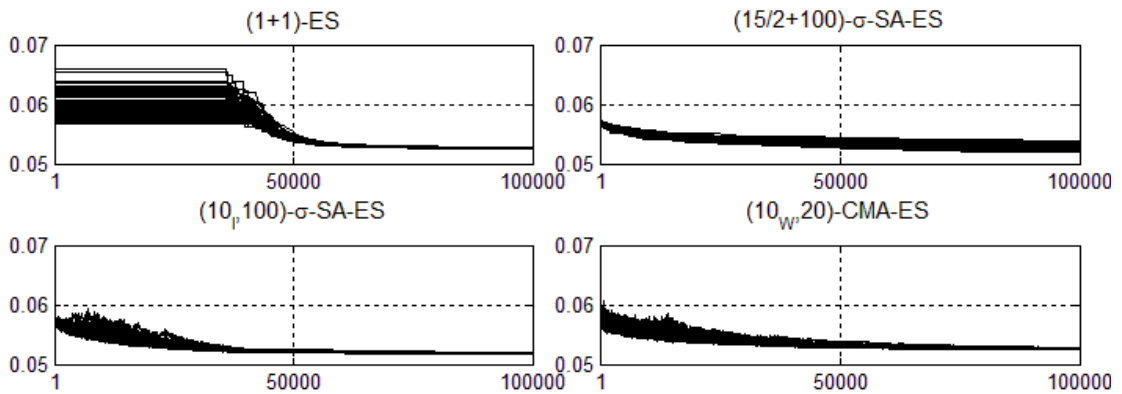


Fig. 5 $\log_{10}(f)$ values of each ES variant vs. objective function evaluations (stationary excitation case, 5% noise corrupted)

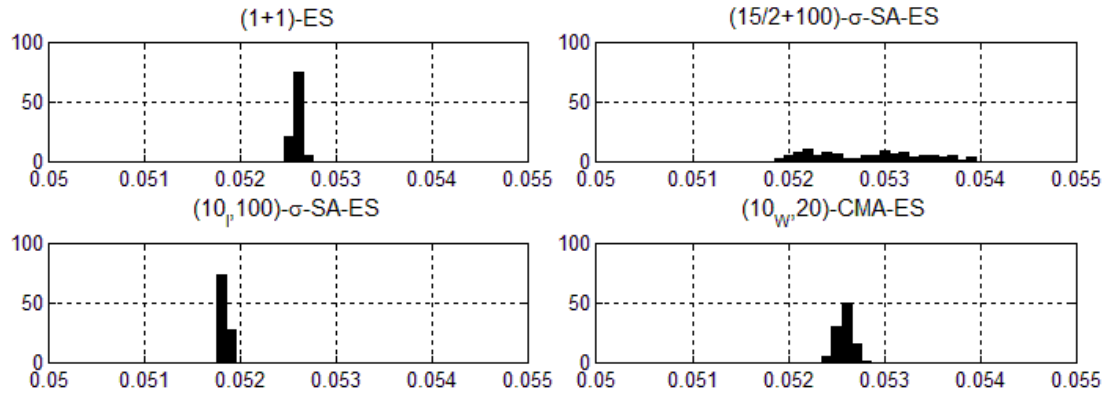


Fig. 6 Histograms of the final $\log_{10}(f)$ values of each ES variant at the end of every Monte Carlo iteration (stationary excitation case, 5% noise corrupted)

The statistical consistency in the noise-corrupted case is significantly improved for the σ -SA-ES variants, as confirmed by all the illustrated data. In addition, the estimates of the vibration modes (Table 10) are again very consistent, diverging slightly from their theoretical counterparts, as the indicated standard deviations imply.

It must be noted though that the case of noise-corrupted vibration response is better handled by using prediction-error state-space models (Verhaegen and Verdult 2007), rather than output-error state-space models, which are used here. Within the context of the current benchmark study, this is an issue that deserves further investigation.

Table 7 Percentage termination reason (stationary excitation case, 5% noise corrupted)

Termination reason	(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)-CMA
Stepsize low bound	0%	0%	0%	0%
Objective function threshold	0%	0%	0%	0%
Max. Function Evaluations	100%	100%	100%	100%

Table 8 Percentage improvement from the subspace estimate (mean value \pm standard deviation, stationary excitation case, 5% noise corrupted)

(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)-CMA
7.75% \pm 0.07%	7.39% \pm 0.97%	9.06% \pm 0.05%	7.77% \pm 0.13%

Table 9 Trace of the MSE of the output normal form parameterization (Eq. (9), stationary excitation case, 5% noise corrupted)

(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)–CMA
5.90×10^{-4}	177×10^{-4}	18.85×10^{-4}	9.66×10^{-4}

Table 10 Estimated structural vibration modes (mean value, stationary excitation case, 5% noise corrupted). Only significant standard deviations are shown

Mode	(1+1)		(15/2+100)		(10 _I ,100)		(10 _w ,20)–CMA	
	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)
1	1.72	5.02	1.72	5.02	1.72	5.02	1.72	5.02
2	5.00	3.17	5.01	3.17	5.01	3.17	5.01	3.18
3	7.84	3.54	7.84	3.54	7.84	3.54	7.84	3.54
4	10.18	4.16	10.18	4.16	10.18	4.16	10.18	4.16
5	12.06	4.64	12.06	4.63	12.06	4.54	12.06	4.63
		± 0.01		± 0.03		± 0.01		± 0.02
6	13.34	5.65	13.33	5.61	13.30	5.37	13.33	5.64
		± 0.04	± 0.02	± 0.12		± 0.03	± 0.01	± 0.06

4.3 Case II: non-stationary excitation

4.3.1 Noise-free estimation

In the non-stationary case the Northridge earthquake is selected to simulate the structure and the estimation is taking place by utilizing 4000 samples (per channel) of excitation–response data. The ES variants are implemented with $\sigma_0 = 10^{-4}$ and $\sigma_t = f_t = 10^{-16}$, while the results are expanded over Figs. 7 and 8 and Tables 11–14, when no noise is added to the data. The following remarks are drawn:

(1+1)–ES: the evolution of the (\log_{10}) objective function vs. evaluation times (Fig. 7) is the same as in Figs. 3 and 5, as again the algorithm starts to descend after approximately 70000

function evaluations, approaching asymptotically the -9.5 line (see also the histogram of Fig. 8). In all the experiments the algorithm terminates due to the objective function evaluation threshold, achieving a mean improvement of approximately 24% (Table 11), in comparison to the subspace estimate. Statistical consistency is verified once more by the histogram of Fig. 7 and the corresponding numerical values of Tables 11-14.

(15/2+100)- σ -SA-ES: better performance is exhibited in this version, compared to the previous one, returning better estimates in all Monte Carlo simulations. From Fig. 7 it follows that iterations are constantly decreasing and that the objective functions lie in the area of -10, as observed also from the histogram of Fig. 8), yet, although at a higher deviation than the one of the (1+1)-ES. The algorithm terminates due to objective function evaluation threshold (Table 11) and achieves a mean improvement of up 35% (Table 12), in comparison to the subspace estimate. The statistical consistency of this ES variant is comparable to the one of (1+1)-ES.

(10_r,100)- σ -SA-ES: as in the stationary case, the comma self-adaptive ES returns the best results in terms of mean objective function values and percentage improvement to the subspace estimate. Figs. 7 and 8 illustrate that the mean value of the best objective function lies in the area of -10.5 with comparable deviation (Table 12). Again experiments are terminated due to objective function threshold and the mean improvement over the subspace estimate is almost 45% (Table 12). Statistical consistency is also comparable to the previous versions, including the estimation of the vibration modes (Table 14).

(10_w,20)-CMA-ES: although it does not return the best objective function values, the CMA-ES returns the best results in terms of statistical consistency and quality of the final estimate, just as in the stationary case. Both Fig. 8 and Table 12 reveal that the dispersion of the final values is the lowest among all ES variants, while the mean value of the best objective function approximates -10, which is the second best. As previously, in all Monte Carlo experiments the algorithm terminates due to objective function evaluation threshold (Table 11) and the mean improvement in comparison to the subspace estimate is about 35% (Table 12).

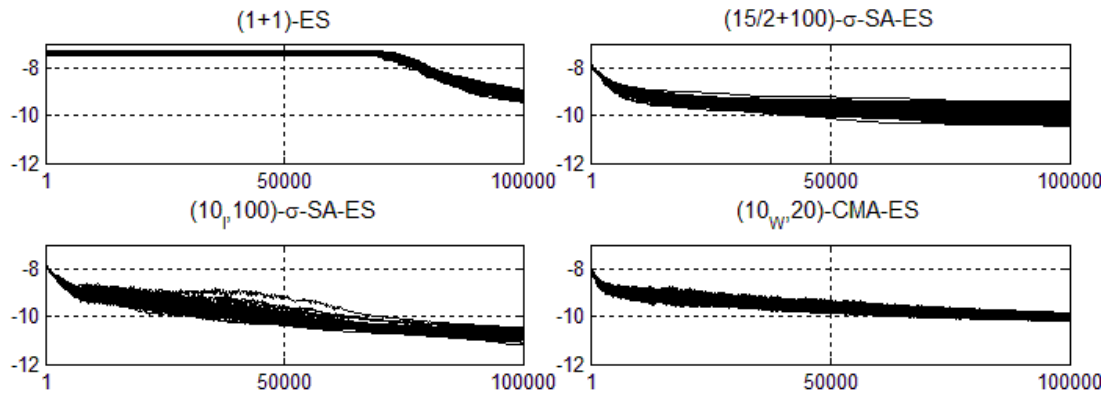


Fig. 7 $\log_{10}(f)$ values of each ES variant vs. objective function evaluations (non-stationary excitation case, noise-free)

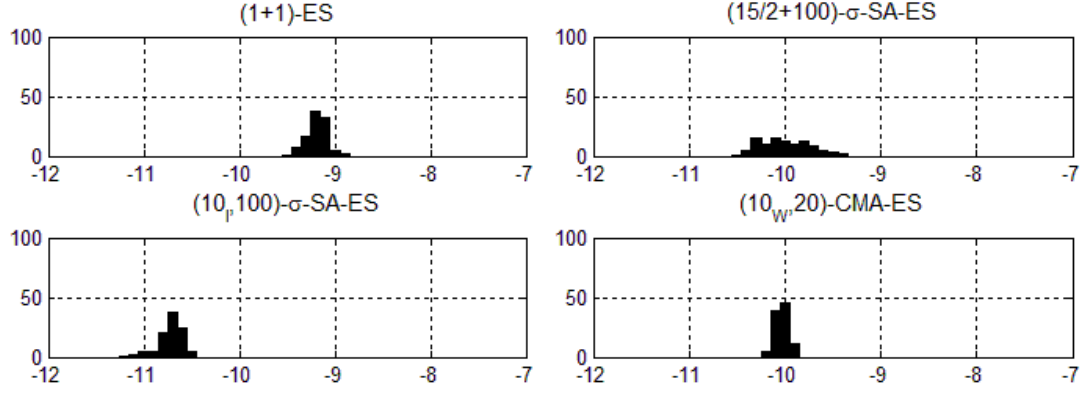


Fig. 8 Histograms of the final $\log_{10}(f)$ values of each ES variant at the end of every Monte Carlo iteration (non-stationary excitation case, noise-free)

Table 11 Percentage termination reason (non-stationary excitation case, noise-free)

Termination reason	(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)–CMA
Stepsize low bound	0%	0%	0%	0%
Objective function threshold	0%	0%	0%	0%
Max. Function Evaluations	100%	100%	100%	100%

Table 12 Percentage improvement from the subspace estimate (mean value \pm standard deviation, non-stationary excitation case, noise-free)

(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)–CMA
23.97% \pm 1.42%	34.95% \pm 3.49%	44.74% \pm 1.79%	35.42% \pm 0.90%

Table 13 Trace of the MSE of the output normal form parameterization (Eq. (9), non-stationary excitation case, noise-free)

(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)–CMA
2.51×10^{-10}	1.22×10^{-10}	5.64×10^{-10}	1.38×10^{-10}

Table 14 Estimated structural vibration modes (mean value, non-stationary excitation case, noise-free). Standard deviations were estimated in the order of 10^{-6} and therefore they're not shown

Mode	(1+1)		(15/2+100)		(10 _p ,100)		(10 _w ,20)-CMA	
	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)
1	1.72	5.00	1.72	5.00	1.72	5.00	1.72	5.00
2	5.00	3.19	5.00	3.19	5.00	3.19	5.00	3.19
3	7.84	3.58	7.84	3.58	7.84	3.58	7.84	3.58
4	10.17	4.13	10.17	4.13	10.17	4.13	10.17	4.13
5	12.06	4.64	12.06	4.64	12.06	4.64	12.06	4.64
6	13.33	5.00	13.33	5.00	13.33	5.00	13.33	5.00

4.3.2 Noise-corrupted estimation

The final Monte Carlo simulation study considered pertains to the noise-corrupted estimation from non-stationary data. To this, the Northridge earthquake again excites the structure and white noise is added to the structural responses of each story at 5% noise-to-signal ratio (standard deviations). The ES variants are implemented with $\sigma_0 = 10^{-4}$ and $\sigma_t = f_t = 10^{-16}$.

The results are illustrated in Figs. 9 and 10 and Tables 15-17. As expected, the type and the quality of the vibration data, in combination to the high dimensionality of the problem, imposed serious difficulties in the estimation process. In all cases the algorithms are terminated due to the maximum objective function evaluation threshold (Table 15), while the mean percentage improvement from the subspace estimate (Table 16) varies from 17.39%, for the (1+1)-ES, to 21.22%, again for the (10_p,100)- σ -SA-ES. Yet, statistical consistency is significantly improved in comparison to the noise-free case (compare Tables 12 and 16) and this includes the estimates of the vibration modes (Table 17). However, the higher modes have not been identified accurately, fact that is justified by the combined effects of the noise corruption and the frequency content of the input. It is also noted that in this case several models resulted unstable from all ES variants, so the statistic of Eq. (9) is not estimated.

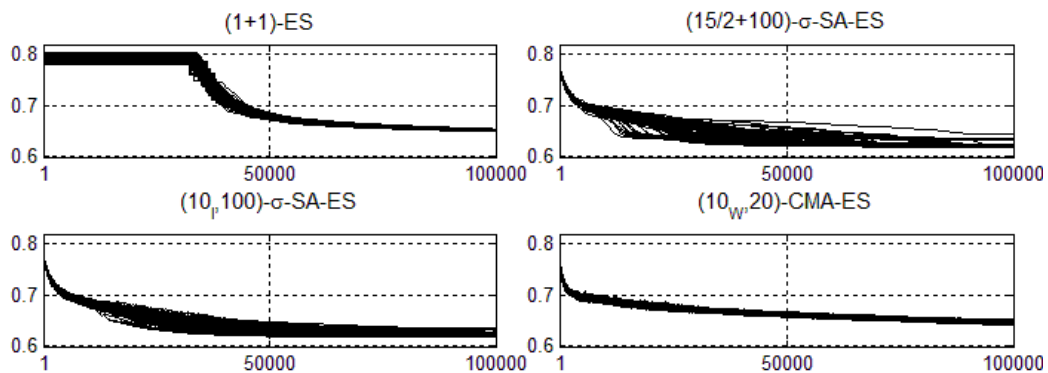


Fig. 9 $\log_{10}(f)$ values of each ES variant vs. objective function evaluations (non-stationary excitation case, 5% noise corrupted)

Table 15 Percentage termination reason (non-stationary excitation case, 5% noise corrupted)

Termination reason	(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)–CMA
Stepsize low bound	0%	0%	0%	0%
Objective function threshold	0%	0%	0%	0%
Max. Function Evaluations	100%	100%	100%	100%

Table 16 Percentage improvement from the subspace estimate (mean value \pm standard deviation, non-stationary excitation case, 5% noise corrupted)

(1+1)	(15/2+100)	(10 _I , 100)	(10 _I , 20)–CMA
17.39% \pm 0.13%	21.12% \pm 0.62%	21.22% \pm 0.56%	17.96% \pm 0.24%

Table 17 Estimated structural vibration modes (mean value, non-stationary excitation case, 5% noise corrupted). Only significant standard deviations are shown

Mode	(1+1)		(15/2+100)		(10 _I ,100)		(10 _w ,20)–CMA	
	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)	f_n (Hz)	ζ_n (%)
1	1.72	5.01	1.72	5.00	1.72	5.00	1.72	5.01
2	5.00	3.12	5.00	3.17	5.00	3.17	5.00	3.13
3	7.84	3.56	7.84	3.58 ± 0.03	7.84	3.59 ± 0.02	7.84 ± 0.01	3.59 ± 0.01
4	10.16 ± 0.02	0.87 ± 1.27	10.11 ± 0.11	1.76 ± 1.58	10.08 ± 0.14	2.70 ± 1.73	10.2 ± 0.02	0.99 ± 1.23
5	10.19	3.50 ± 1.25	10.19 ± 0.06	3.51 ± 1.30	12.26 ± 0.25	4.32 ± 3.07	10.19	3.51 ± 1.20
6	12.09 ± 0.01	4.50 ± 0.11	12.08 ± 0.02	4.86 ± 0.31	12.07 ± 0.02	5.13 ± 0.29	12.1 ± 0.02	4.55 ± 0.14

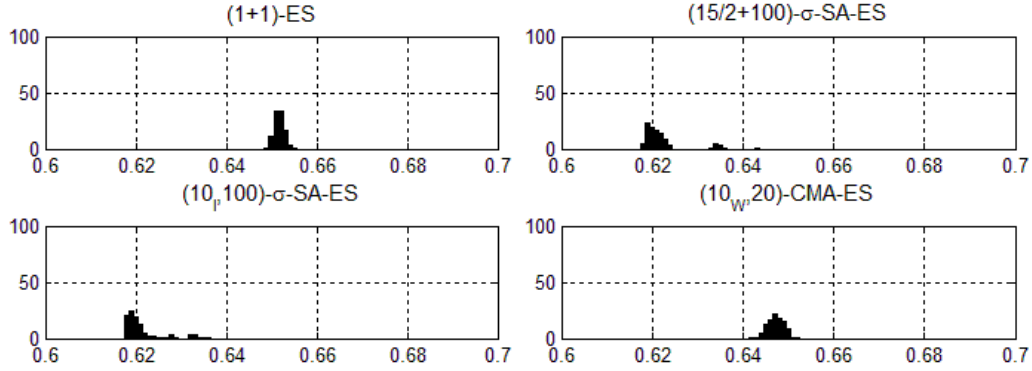


Fig. 10 Histograms of the final $\log_{10}(f)$ values of each ES variant at the end of every Monte Carlo iteration (non-stationary excitation case, 5% noise corrupted)

5. Conclusions

The aim of this benchmark study was to investigate the applicability of stochastic optimization theory to the structural identification problem using state-space model representations. To this, four widely accepted variants of ES were utilized and a demanding structural estimation problem was considered, especially in terms of dimensionality. The results of the Monte Carlo simulation showed that all variants are capable of handling the state-space estimation problem, at the cost of increased computational complexity. Regarding the individual ES variants, the following remarks are made:

(1+1)-ES: in all simulation cases the algorithm stuck for a significant amount of iterations and then started to descend. Mean improvement from the subspace estimates varied from 7% up to 28%, while statistical consistency (small deviations) was very good. Modal parameters were accurately estimated, except from the last case, at which biased higher modes were resulted.

(15/2+100)-σ-SA-ES: the plus version of the SA-ES showed better performance than the one of (1+1)-ES, exhibiting a constant decrease in the objective function value through the whole spectrum of iterations, and producing better final values and mean improvements from the subspace estimates that varied from 7% up to 48%. Yet, the statistical consistency was not as good as the one of (1+1)-ES, especially in the first Monte Carlo simulation study, which is an order of magnitude higher. Modal parameters were accurately estimated, except from the last case, at which biased higher modes were resulted.

(10I, 100)-σ-SA-ES: the comma version of the SA-ES presented similar performance to its plus counterpart. However, this ES variant produced the best results in terms of final objective function values and mean improvements from the subspace estimates, as the latter varied from 9% to 61%. Statistical consistency was comparable to the plus version as well, while modal parameters were accurately estimated: even in the last simulation case, the bias of the higher modes is the lowest.

(10w, 20)-CMA-ES: this ES variant provided indications of “compromise” between final values and statistical consistency. Thus, while it did not produce the extreme final values of (10I,100)-σ-SA-ES, the best objective function values were high, while, at the same time,

statistical consistency was kept to the levels of the (1+1)–ES variant (in fact it was better in half the cases). The mean improvement from the subspace estimates varied from 8% up to 34% and the quality of the modal parameters' estimates was, more or less, the same as before.

The encouraging results suggest further research towards the stochastic structural identification problem via state–space models. Decrease in the computational complexity and introduction of prediction–error terms, in order to tackle with noise–corrupted data, is under ongoing research. In addition, investigation of the role, as well as corresponding adaptation of the exogenous ES parameters to the specific requirements of this model–based structural identification problem is also being considered.

Acknowledgments

The author would like to acknowledge the European Commission for funding SmartEN (Grant No. 238726) under the Marie Curie ITN FP7 program, as the research work presented here is supported by this program. The opinions expressed in this paper do not necessarily reflect those of the sponsors.

References

- Arnold, D.V. (2006), “Weighted multirecombination evolution strategies”, *Theor. Comput. Sci.*, **361**(1), 18-37.
- Arnold, D.V. and Beyer, H.G. (2002), “Performance analysis of evolution strategies with multi-recombination in high-dimensional RN-search spaces disturbed by noise”, *Theor. Comput. Sci.*, **289**(1), 629-647.
- Auger, A. (2009), “Benchmarking the (1+1) evolution strategy with one-Fifth success rule on the BBOB-2009 function testbed”, *Proceedings of the GECCO'09*, Montréal Québec, Canada.
- Beyer, H.G. and Sendhoff, B. (2008), *Covariance matrix adaptation revisited—the CMA evolution strategy*, (Eds., Rudolph, G., Jansen, Th., Lucas, S.M., Poloni, C. and Beume, N.), Parallel Problem Solving from Nature - PPSN X, Springer.
- Beyer, H.G. and Schwefel, H.P. (2002), “Evolution strategies: a comprehensive introduction”, *Nat. Comput.*, **1**(1), 3-52.
- Casciati, S. (2008), “Stiffness identification and damage localization via differential evolution algorithms”, *Struct. Control Health Monit.*, **15**(3), 436-449.
- Casciati, S. (2010), “Statistical approach to a SHM benchmark problem”, *Smart Struct. Syst.*, **6**(1), 17-27.
- Dertimanis, V.K., Koulocheris, D., Vrazopoulos, H. and Kanarachos, A. (2003), “Time-series parametric modeling using Evolution Strategy with deterministic mutation operators”, *Proceedings of the International Conference on Intelligent Control Systems and Signal Processing*, Faro, Portugal.
- Faravelli, L. and Casciati, S. (2004) “Structural damage detection and localization by response change diagnosis”, *Struct. Saf.*, **6**(2), 104-115.
- Fassois, S.D. (2001), “MIMO LMS-ARMAX identification of vibrating structures-Part I: the method”, *Mech. Syst. Signal Pr.*, **15**(4), 737-758.
- Franco, G., Betti, R. and Luş, H. (2004), “Identification of structural systems using an evolutionary strategy”, *J. Eng. Mech. -ASCE*, **130**(10), 1125-1139.
- Hansen, N. (2006), *The CMA evolution strategy: a comparing review*, (Eds., J.A. Lozano, P. Larrañga, I. Inza and E. Bengoetxea), Towards a new evolutionary computation: Advances in estimation of distribution algorithms, Springer.
- Hansen, N. and Ostermeier, A. (2001), “Completely derandomized self-adaptation in evolution strategies”,

- Evol. Comput.*, **9**(2), 159-195.
- Hansen, N. and Ostermeier, A. (1997), "Convergence properties of evolution strategies with the derandomized covariance matrix adaptation: the $(\mu/\mu_i, \lambda)$ -CMA-ES", *Proceedings of the 5th European Congress on Intelligent techniques and Soft Computing*, Aachen, Germany.
- Huang, C.S. (2001), "Structural identification from ambient vibration using the multivariate AR model", *J. Sound Vib.*, **241**(3), 337-359.
- Huang, J.N. and Pappa, R.S. (1985), "An eigensystem realization algorithm (ERA) for modal parameter identification and model reduction", *J. Guid. Control Dynam.*, **8**, 620-627.
- Katayama, T. (2005), *Subspace methods for system identification*, Springer-Verlag, Berlin, Germany.
- Koulocheris, D., Dertimanis, V.K. and Spentzas, C.N. (2008), "Parametric identification of vehicle structural characteristics", *Forsch. Ingenieurwes.*, **68**(4), 173-181.
- Koulocheris, D., Dertimanis, V.K. and Vrazopoulos, H. (2004), "Evolutionary parametric identification of dynamic systems", *Forsch. Ingenieurwes.*, **72**(1), 39-51.
- Koulocheris, D., Dertimanis, V.K. and Vrazopoulos, H. (2003), "Vehicle suspension system identification using evolutionary algorithms", *Proceedings of the EUROGEN 2003*, Barcelona, Spain.
- Kruisselbrink, J.W. (2012), *Evolution strategies for robust optimization*, PhD Thesis, Leiden University, Leiden, The Netherlands.
- Landau, I.D. and Zito, G. (2006), *Digital control systems: design, identification and implementation*, Springer-Verlag, London, UK.
- Ljung, L. (1999), *System identification: theory for the user*, 2nd Ed., Prentice-Hall Inc., Englewood Cliffs, NJ, USA.
- Mc Kelvey, T. and Helmersson, A. (1997), "System identification using an over-parameterized model class - Improving the optimization algorithm", *Proceedings of the 36th IEEE Conference on Decision and Control*, San Diego, CA.
- Ostermeier, A., Gawelczyk, A. and Hansen, N. (1994), "Step-size adaptation based on non-local use of selection information", in *Parallel Problem Solving from Nature — PPSN III*, Springer.
- Tang, H.S., Xue S.T. and Fan, C. (2008), "Differential evolution strategy for structural system identification", *Comput. Struct.*, **86**(21-22), 2004-2012.
- Van Overschee, P. and De Moor, B. (1996), *Subspace identification for linear systems: theory - Implementation - Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Verhaegen, M. and Dewilde, P. (1992), "Subspace model identification Part 1. the output-error state-space model identification class of algorithms", *Int. J. Control*, **56**(5), 1187-1210.
- Verhaegen, M. and Verdult, V. (2007), *Filtering and system identification: a least squares approach*, 2nd Ed., Prentice-Hall Inc., Englewood Cliffs, NJ, USA.
- Viberg, M. (1995), "Subspace-based methods for the identification of linear time-invariant systems", *Automatica*, **31**(12), 1835-1851.
- Zhang, Z., Koh, C.G. and Duan, W.H. (2010a), "Uniformly sampled genetic algorithm with gradient search for structural identification - Part I: Global search", *Comput. Struct.*, **88**(15-16), 949-962.
- Zhang, Z., Koh, C.G. and Duan, W.H. (2010b), "Uniformly sampled genetic algorithm with gradient search for structural identification - Part II: Local search", *Comput. Struct.*, **88**(19-20), 1149-1161.