

GEOPHYSICS®

Particle Swarm Optimization of 2D Magnetotelluric data

| | |
|--------------------|--|
| Journal: | <i>Geophysics</i> |
| Manuscript ID | GEO-2018-0166.R1 |
| Manuscript Type: | Technical Paper |
| Keywords: | 2D, algorithm, electromagnetics, magnetotelluric, optimization |
| Area of Expertise: | Electrical and Electromagnetic Methods, Engineering and Environmental Geophysics |
| | |

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

GEOPHYSICS SAMPLE TITLE PAGE

Particle Swarm Optimization of 2D Magnetotelluric data

Francesca Pace¹, Alessandro Santilano², and Alberto Godio¹

Right Running Head: PSO of 2D MT data

1: Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di
Torino, Italy. E-mail: francesca.pace@polito.it; alberto.godio@polito.it.

2: Institute of Geosciences and Earth Resources- National Research Council (IGG-CNR), Pisa,
Italy. E-mail: alessandro.santilano@igg.cnr.it

PARTICLE SWARM OPTIMIZATION OF 2D MAGNETOTELLURIC DATA

ABSTRACT

We implement the Particle Swarm Optimization (PSO) algorithm for the two-dimensional (2D) magnetotelluric (MT) inverse problem. We first validate PSO on two synthetic models of different complexity and then apply it to an MT benchmark for real field data, the COPROD2 data set (Canada). We pay particular attention to the selection of the PSO input parameters to properly address the complexity of the 2D MT inverse problem. We enhance the stability and convergence of the solution of the geophysical problem by applying the hierarchical PSO with time-varying acceleration coefficients (HPSO-TVAC). Moreover, we parallelize the code to reduce the computation time, since PSO is a computationally demanding global search algorithm. The inverse problem was solved for the synthetic data both by giving a priori information at the beginning and by using a random initialization. The a priori information was given to a small number of particles as initial position within the search space of solutions, so that the swarming behavior was only slightly influenced. We demonstrate that there is no need for the a priori initialization to obtain robust 2D models, since results are largely comparable with results from randomly initialized PSO. The optimization of the COPROD2 data set provides a resistivity model of the Earth in line with results from previous interpretations. Our results suggest that the 2D MT inverse problem can be successfully addressed by means of computational swarm intelligence.

INTRODUCTION

The interpretation of geophysical data requires the solution of the inverse problem, which is, in most cases, nonlinear and ill-posed. During the past three decades, global search algorithms as inversion methods have become of growing interest, since the probabilistic or evolutionary approach has been adopted to find the optimum solution, which is affected by non-uniqueness. The most important global-search algorithms generally used for the inversion of geophysical data are simulated

1
2
3 38 annealing (SA), genetic algorithm (GA) (Sen and Stoffa, 2013), ant-colony algorithm (ACO) (Yuan
4
5 39 et al., 2009), and particle swarm optimization (PSO) (Shaw and Srivastava, 2007).
6
7
8 40
9
10
11 41 The inversion of magnetotelluric (MT) data is usually based on algorithms such as Occam,
12
13 42 Non-Linear Conjugate Gradient (NLCG) and Gauss-Newton (GN), which are now widely recognized
14
15 43 as milestones among two-dimensional (2D) and three-dimensional (3D) MT inversion codes
16
17 44 (Avdeev, 2005; Boerner, 2010; Siripunvaraporn, 2012; Ghaedrahmati et al., 2014; Newman, 2014).
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
5510
5511
5512
5513
5514
5515
5516
5517
5518
5519
5520
5521
5522
5523
5524
5525
5526
5527
5528
5529
5530
5531
5532
5533
5534
5535
5536
5537
5538
5539
5540
5541
5542
5543
5544
5545
5546
5547
5548
5549
5550
5551
5552
5553
5554
5555
5556
5557
5558
5559
55510
55511
55512
55513
55514
55515
55516
55517
55518
55519
55520
55521
55522
55523
55524
55525
55526
55527
55528
55529
55530
55531
55532
55533
55534
55535
55536
55537
55538
55539
55540
55541
55542
55543
55544
55545
55546
55547
55548
55549
55550
55551
55552
55553
55554
55555
55556
55557
55558
55559
555510
555511
555512
555513
555514
555515
555516
555517
555518
555519
555520
555521
555522
555523
555524
555525
555526
555527
555528
555529
555530
555531
555532
555533
555534
555535
555536
555537
555538
555539
555540
555541
555542
555543
555544
555545
555546
555547
555548
555549
555550
555551
555552
555553
555554
555555
555556
555557
555558
555559
5555510
5555511
5555512
5555513
5555514
5555515
5555516
5555517
5555518
5555519
5555520
5555521
5555522
5555523
5555524
5555525
5555526
5555527
5555528
5555529
5555530
5555531
5555532
5555533
5555534
5555535
5555536
5555537
5555538
5555539
5555540
5555541
5555542
5555543
5555544
5555545
5555546
5555547
5555548
5555549
5555550
5555551
5555552
5555553
5555554
5555555
5555556
5555557
5555558
5555559
55555510
55555511
55555512
55555513
55555514
55555515
55555516
55555517
55555518
55555519
55555520
55555521
55555522
55555523
55555524
55555525
55555526
55555527
55555528
55555529
55555530
55555531
55555532
55555533
55555534
55555535
55555536
55555537
55555538
55555539
55555540
55555541
55555542
55555543
55555544
55555545
55555546
55555547
55555548
55555549
55555550
55555551
55555552
55555553
55555554
55555555
55555556
55555557
55555558
55555559
555555510
555555511
555555512
555555513
555555514
555555515
555555516
555555517
555555518
555555519
555555520
555555521
555555522
555555523
555555524
555555525
555555526
555555527
555555528
555555529
555555530
555555531
555555532
555555533
555555534
555555535
555555536
555555537
555555538
555555539
555555540
555555541
555555542
555555543
555555544
555555545
555555546
555555547
555555548
555555549
555555550
555555551
555555552
555555553
555555554
555555555
555555556
555555557
555555558
555555559
5555555510
5555555511
5555555512
5555555513
5555555514
5555555515
5555555516
5555555517
5555555518
5555555519
5555555520
5555555521
5555555522
5555555523
5555555524
5555555525
5555555526
5555555527
5555555528
5555555529
5555555530
5555555531
5555555532
5555555533
5555555534
5555555535
5555555536
5555555537
5555555538
5555555539
5555555540
5555555541
5555555542
5555555543
5555555544
5555555545
5555555546
5555555547
5555555548
5555555549
5555555550
5555555551
5555555552
5555555553
5555555554
5555555555
5555555556
5555555557
5555555558
5555555559
55555555510
55555555511
55555555512
55555555513
55555555514
55555555515
55555555516
55555555517
55555555518
55555555519
55555555520
55555555521
55555555522
55555555523
55555555524
55555555525
55555555526
55555555527
55555555528
55555555529
55555555530
55555555531
55555555532
55555555533
55555555534
55555555535
55555555536
55555555537
55555555538
55555555539
55555555540
55555555541
55555555542
55555555543
55555555544
55555555545
55555555546
55555555547
55555555548
55555555549
55555555550
55555555551
55555555552
55555555553
55555555554
55555555555
55555555556
55555555557
55555555558
55555555559
555555555510
555555555511
555555555512
555555555513
555555555514
555555555515
555555555516
555555555517
555555555518
555555555519
555555555520
555555555521
555555555522
555555555523
555555555524
555555555525
555555555526
555555555527
555555555528
555555555529
555555555530
555555555531
555555555532
555555555533
555555555534
555555555535
555555555536
555555555537
555555555538
555555555539
555555555540
555555555541
555555555542
555555555543
555555555544
555555555545
555555555546
555555555547
555555555548
555555555549
555555555550
555555555551
555555555552
555555555553
555555555554
555555555555
555555555556
555555555557
555555555558
555555555559
5555555555510
5555555555511
5555555555512
5555555555513
5555555555514
5555555555515
5555555555516
5555555555517
5555555555518
5555555555519
5555555555520
5555555555521
5555555555522
5555555555523
5555555555524
5555555555525
5555555555526
5555555555527
5555555555528
5555555555529
5555555555530
5555555555531
5555555555532
5555555555533
5555555555534
5555555555535
5555555555536
5555555555537
5555555555538
5555555555539
5555555555540
5555555555541
5555555555542
5555555555543
5555555555544
5555555555545
5555555555546
5555555555547
5555555555548
5555555555549
5555555555550
5555555555551
5555555555552
5555555555553
5555555555554
5555555555555
5555555555556
5555555555557
5555555555558
5555555555559
55555555555510
55555555555511
55555555555512
55555555555513
55555555555514
55555555555515
55555555555516
55555555555517
55555555555518
55555555555519
55555555555520
55555555555521
55555555555522
55555555555523
55555555555524
55555555555525
55555555555526
55555555555527
55555555555528
55555555555529
55555555555530
55555555555531
55555555555532
55555555555533
55555555555534
55555555555535
55555555555536
55555555555537
55555555555538
55555555555539
55555555555540
55555555555541
55555555555542
55555555555543
55555555555544
55555555555545
55555555555546
55555555555547
55555555555548
55555555555549
55555555555550
55555555555551
55555555555552
55555555555553
55555555555554
55555555555555
55555555555556
55555555555557
55555555555558
55555555555559
555555555555510
555555555555511
555555555555512
555555555555513
555555555555514
555555555555515
555555555555516
555555555555517
555555555555518
555555555555519
555555555555520
555555555555521
555555555555522
555555555555523
555555555555524
555555555555525
555555555555526
555555555555527
555555555555528
555555555555529
555555555555530
555555555555531
555555555555532
555555555555533
555555555555534
555555555555535
555555555555536
555555555555537
555555555555538
555555555555539
555555555555540
555555555555541
555555555555542
555555555555543
555555555555544
555555555555545
555555555555546
555555555555547
555555555555548
555555555555549
555555555555550
555555555555551
555555555555552
555555555555553
555555555555554
555555555555555
555555555555556
555555555555557
555555555555558
555555555555559
5555555555555510
5555555555555511
5555555555555512
5555555555555513
5555555555555514
5555555555555515
5555555555555516
5555555555555517
5555555555555518
5555555555555519
5555555555555520
5555555555555521
5555555555555522
5555555555555523
5555555555555524
5555555555555525
5555555555555526
5555555555555527
5555555555555528
5555555555555529
5555555555555530
5555555555555531
5555555555555532
5555555555555533
5555555555555534
5555555555555535
5555555555555536
5555555555555537
5555555555555538
5555555555555539
5555555555555540
5555555555555541
5555555555555542
5555555555555543
5555555555555544
5555555555555545
5555555555555546
5555555555555547
5555555555555548
5555555555555549
5555555555555550
5555555555555551
5555555555555552
5555555555555553
5555555555555554
5555555555555555
5555555555555556
5555555555555557
5555555555555558
5555555555555559
55555555555555510
55555555555555511
55555555555555512
55555555555555513
55555555555555514
55555555555555515
55555555555555516
55555555555555517
55555555555555518
55555555555555519
55555555555555520
55555555555555521
55555555555555522
55555555555555523
55555555555555524
55555555555555525
55555555555555526
55555555555555527
55555555555555528
55555555555555529
55555555555555

1
2
3 includes the interpretation of Vertical Electrical Sounding (Fernández Martínez et al., 2010), gravity
4
5 (Darisma et al., 2017) and Multi-Transient electromagnetic data (Olalekan and Di, 2017). The main
6
7 limitation of these works is that they analyze either one-dimensional field data or over-simplistic 2D
8
9 synthetic models. Other scientific applications of the PSO algorithm are artificial neural networks,
10
11 biomedical engineering, hydrogeology, electronics, electromagnetics, power systems, robotics, and
12
13 signal processing (Adhan and Bansal, 2017; Poli, 2008 and references therein).
14
15

The present paper focuses on the implementation of the PSO algorithm for the 2D MT inverse problem. A preliminary application of this method to MT and Audio-MT synthetic data has been presented in Pace et al. (2017). The novelty of this paper concerns the validation of the method on two MT synthetic models of different complexity and, for the first time to the authors' knowledge, the application to real field data, the COPROD2 data set (Jones, 1993a). This data set was made available to the electromagnetic-induction scientific community with the aim of comparing different techniques for 2D MT inversion (Jones, 1993b and references therein). Since several inversion solutions have been made available so far, COPROD2 represents an interesting (and challenging) field data set for the application of our method. We started from the PSO code of Ebbesen et al. (2012), but then we modified that generic MATLAB code for our specific purpose. The efficiency of the PSO algorithm was improved by applying the principle of hierarchical PSO with time-varying acceleration coefficients (HPSO-TVAC) (Ratnaweera et al., 2004). Previous works on PSO applied to the geophysical inverse problem have always considered constant values for both social and cognitive accelerations of particles (Shaw and Srivastava, 2007; Godio and Santilano, 2018; Santilano et al., 2018). However, this assumption is not adequate for the 2D inverse problem due to its high-dimensionality and complex searching behavior. We carried out a detailed sensitivity analysis to find the most appropriate values of the time-varying accelerations: their iterative variation improved the convergence speed of the algorithm and prevented the solution from being trapped in some local minima. The tuning of the social and cognitive accelerations of the particles was hence crucial to

1
2
3 88 finally achieve the convergence of the solution. In addition, a new parallelized version of the code
4
5 89 was developed with the aim of overcoming the time-consuming nature of PSO, which is
6
7 90 computationally demanding, like the other global search algorithms. We modified the released PSO
8
9 code to be run on a High Performance Computing (HPC) cluster.
10
11
12

13 92 The ensuing sections are organized as follows: First, we explain the way the swarm
14
15 93 intelligence principle is applied to the geophysical inverse problem and, specifically, to the 2D MT
16
17 94 inverse problem. Then, the first synthetic model is adopted to show the calibration of the main tuning
18
19 95 parameters of PSO, the accelerations, and population size. Once the input arguments are set, the two
20
21 96 different synthetic models of MT data are optimized and the obtained results are illustrated. After the
22
23 97 method validation with the synthetic models, the PSO algorithm is applied to the COPROD2 data set,
24
25
26 98 and the final resistivity model is compared with results reported in the literature and using well-
27
28 99 established algorithms. Finally, the computational improvements are outlined.

100 PSO APPLICATION TO 2D MAGNETOTELLURIC INVERSE PROBLEM

101 The PSO algorithm is a population-based algorithm that simulates the self-organizing
102 behavior of species living in groups, such as flocks of birds or schools of fish. The way they share
103 knowledge to search for food or find the best reciprocal distance in motion fascinated Kennedy and
104 Eberhart (1995) so strongly that they proposed applying this evolutionary approach to the
105 optimization of nonlinear problems. Simple interactions between individuals yield a complex
106 collective behavior, meaning that each individual is able to adapt and derive new and coherent
107 behaviors in case of changes in the external environment. The most striking feature of this method is
108 that every particle has a memory component that rules its behavior. This is influenced by both the
109 cognitive knowledge of the particle and the experience of its neighbors, whose leadership can be
110 emulated. Pivotal references for computational swarm intelligence are Kennedy et al. (2001) and
111 Engelbrecht (2007).

The basic concept of PSO application to geophysics is that each particle of the swarm represents a possible solution of the MT inverse problem, that is, an electrical-resistivity model. Since the solution of the problem is affected by non-uniqueness, the search space of solutions needs to be fully explored in order to find the best model which fits the observed data. This need is fulfilled by the adaptive and swarming behavior of the particles. During the optimization process, iteration after iteration, the particles “fly” within the search space, bounded between a minimum and a maximum resistivity value. At the end of the swarming, the optimized solution is identified. Readers can find a detailed description of the application of PSO to the geophysical 1D inverse problem in Fernández Martínez et al. (2010a, 2010b) for VES and self-potential methods and Santilano et al. (2018) for MT.

Since the implementation of the PSO algorithm for the 2D MT problem required a high number of particles forming the swarm and numerous iterations to achieve convergence, the standard release of the code for MATLAB appeared to need some modifications. Several variations of the PSO algorithm have been proposed in order to accelerate convergence and avoid a solution trapped in local minima (Zhan et al., 2009). The PSO variant that showed improved outcomes, with respect to the standard PSO, was the hierarchical PSO with time-varying acceleration coefficients (HPSO-TVAC) (Ratnaweera et al., 2004). This method takes the social and cognitive behavior of particles into account to enhance the solution convergence and stability.

In the 2D MT problem, the particle of the swarm represents a resistivity model, which is a vector whose elements are the resistivity values of the 2D mesh cells. Each particle of the swarm changes its position \mathbf{x} within the search space by means of the velocity vector \mathbf{v} . The vectors \mathbf{x} and \mathbf{v} are updated iteration by iteration according to

$$\mathbf{v}_i^{k+1} = \omega^k \mathbf{v}_i^k + \alpha_1^k \gamma_1 (\mathbf{P}_i - \mathbf{x}_i^k) + \alpha_2^k \gamma_2 (\mathbf{G} - \mathbf{x}_i^k), \quad (1)$$

$$\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \mathbf{v}_i^{k+1}, \quad (2)$$

1
 2
 3 135 where: $i = [1, \dots, N]$, N is the number of particles forming the swarm; k is the current iteration number;
 4
 5 136 \mathbf{x}_i^k and \mathbf{v}_i^k are the current vectors of position and velocity of the i^{th} particle, respectively; ω^k is the
 6
 7 137 inertia weight that linearly decreases from 0.9 (first iteration) to 0.4 (last iteration) in order to balance
 8
 9 138 the momentum remembered from the previous iteration (Shi and Eberhart, 1998); α_1^k is the cognitive
 10
 11 139 acceleration towards the best particle position \mathbf{P} , also called “local best”; α_2^k is the social acceleration
 12
 13 140 towards the best global position \mathbf{G} (or “global best”) found by the group leader; and γ_1 and $\gamma_2 \in [0, 1]$
 14
 15 141 are uniformly distributed random values which provide stochastic perturbation. At the beginning
 16
 17 142 ($k=0$), the velocity vector (\mathbf{v}_i^0) is zero and the position vector (\mathbf{x}_i^0) is randomly initialized. Then (k
 18
 19 143 >0), the particle velocity (\mathbf{v}_i^k) changes according to three terms: cognitive memory α_1^k , social
 20
 21 144 attraction α_2^k , and inertia component ω^k . Finally, the particle position \mathbf{x}_i^k is updated. While all the
 22
 23 145 previous PSO applications to the inverse problem adopted constant accelerations (Fernández
 24
 25 146 Martínez et al., 2010a; Godio and Santilano, 2018; Santilano et al., 2018), we set both acceleration
 26
 27 147 coefficients to vary at each iteration, according to the HPSO-TVAC approach. At the beginning of
 28
 29 148 the HPSO-TVAC optimization, α_1 was larger than α_2 , then they linearly reversed. In this way, at the
 30
 31 149 start the diversity of the swarm ensured the search space exploration (high α_1^k), and, at the end, the
 32
 33 150 exploitation of the best regions and the convergence towards the best solution were enabled (high
 34
 35 151 α_2^k). The resulting adaptive behavior was hence enhanced. In more detail, the cognitive and social
 36
 37 152 accelerations changed according to, respectively,

$$\alpha_1^k = \alpha_1^{\max} - (\alpha_1^{\max} - \alpha_1^{\min}) \left(\frac{k-1}{\max(k)-1} \right), \quad (3)$$

$$\alpha_2^k = \alpha_2^{\min} + (\alpha_2^{\max} - \alpha_2^{\min}) \left(\frac{k-1}{\max(k)-1} \right), \quad (4)$$

53 155 where: α^k is the acceleration value at the current iteration k ; α_1^{\max} and α_2^{\max} are the maximum
 54
 55 156 acceleration values for the cognitive and social accelerations, respectively; α_1^{\min} and α_2^{\min} are the
 56
 57 157 minimum acceleration values for the cognitive and social accelerations, respectively; and $\max(k)$ is
 58
 59 158 the maximum number of iterations set for the optimization (Engelbrecht, 2007, p. 313, and references

1
2
3 159 therein). So, at the first iteration ($k=1$), $\alpha_1^k = 1 = \alpha_1^{\max}$ and $\alpha_2^k = 1 = \alpha_2^{\min}$, while, at the last iteration (k
4
5 160 $= \max(k)$), $\alpha_1^k = \max(k) = \alpha_1^{\min}$ and $\alpha_2^k = \max(k) = \alpha_2^{\max}$. In addition, the accelerations were chosen to
6
7 161 satisfy the stability solution conditions (Perez and Behdinan, 2007):
8
9
10
11 162 $\alpha_1 + \alpha_2 < 4,$ (5)
12
13
14
15 163 $\frac{\alpha_1 + \alpha_2}{2} - 1 < \omega < 1.$ (6)
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The values of the accelerations influenced the way the particles explored the model space and changed their trajectory with respect to the local and global bests. A thorough sensitivity analysis on the PSO control parameters, both accelerations and inertia, can be found in Ratnaweera et al. (2004) and in Fernández Martínez al. (2010a; 2010b). These works tested several benchmark functions and identified the best ranges of these values, ensuring the convergence and stability of the PSO algorithm. Starting from their results, and obeying equations 5 and 6, we performed some tests to assess the influence of several acceleration values on the solution of the MT inverse problem. For both cognitive and social accelerations, we adopted three different maximum values, α_1^{\max} and α_2^{\max} equal to 1.5, 2 and 2.75, and three different minimum values, α_1^{\min} and α_2^{\min} equal to 0.25, 0.5 and 0.75. This sensitivity analysis was applied to synthetic model 1 and the results are presented in the next section.

175 The fitness function

The final goal of the optimization process is the minimization of the fitness or objective function. The particle with the lowest fitness value is awarded with the global best position G and is going to attract neighbors depending on the social acceleration α_2 . The fitness function we adopted was the same as that of Everett and Schultz (1993) for the calculation of the data misfit, while the Occam-like regularization was added as proposed by deGroot-Hedlin and Constable (1990). Therefore, for 2D MT data, the function to be minimized was

$$F(\mathbf{m}) = \left(\frac{1}{M} \left\| \frac{\log(\rho_{a,o}) - \log(\rho_{a,p})}{\Delta\rho_{a,o}} \right\|_2^2 + \frac{1}{M} \left\| \frac{\Phi_o - \Phi_p}{\Delta\Phi_o} \right\|_2^2 \right)^{1/2} + \lambda_x \|\partial_x \mathbf{m}\|_2 + \lambda_z \|\partial_z \mathbf{m}\|_2, \quad (7)$$

where: $\rho_{a,o}$ and $\rho_{a,p}$ are observed and predicted apparent resistivity, respectively; Φ_o and Φ_p are observed and predicted impedance phases, respectively; $\Delta\rho_{a,o}$ and $\Delta\Phi_o$ are the errors in observed apparent resistivity and phase, respectively; M is the number of degrees of freedom, i.e., the number of evaluated data; λ_x and λ_z are the Lagrange multipliers in the x- and z-direction, respectively, set as the tradeoff between the model and data misfit to regulate the model roughness, and $\partial_x \mathbf{m}$ and $\partial_z \mathbf{m}$ are the first derivatives of the model solution along the x- and z-directions, respectively. The solution \mathbf{m} is the electrical-resistivity model, i.e., the vector of resistivity values of the 2D domain. This vector has as many elements as the grid cells of the 2D mesh and is represented by the particles of the swarm.

At each iteration, the particle which best minimizes the objective function is assumed as the global best solution (\mathbf{G}), while the other particles can be either attracted or driven away looking for other solutions in the search space. At the end of the optimization, the particle with the minimum $F(\mathbf{m})$ is selected as the final solution and the majority of the other particles converge to it (swarming behavior). Apparent-resistivity values were transformed to their logarithmic values since they can cover different orders of magnitude. The first part in the right-hand side of equation 7 addresses the minimization between observed data, apparent resistivity (ρ_a) and impedance phase (Φ), and predicted data computed by the forward modeling. This calculation of the misfit is defined as the square root of the sum of two squared Euclidean norms, since ρ_a and Φ can have different orders of magnitude and ranges. The forward modeling incorporates the physics of the problem and, starting from the assumed model \mathbf{m} , predicts the responses ρ_a and Φ for each particle of the swarm. The remaining part of equation 7 was added in order to minimize the roughness of the model solution \mathbf{m} : In both horizontal and vertical directions, the differencing operator on \mathbf{m} was weighted by the Lagrange multiplier λ . This approach is the Occam-like optimization and has been adopted for 1D MT problem in Godio and Santilano (2018) and Santilano et al. (2018). In this way, the minimization

1
2
3 206 of the objective function looks for the smoothest model that fits the data, thus ensuring a balance
4
5 207 between the data fitting and the roughness of the model. The value of λ was appropriately chosen
6
7 208 following the L-curve criterion (Farquharson and Oldenburg, 2004). It consists in finding the optimal
8
9 209 tradeoff between the misfit of the data and the roughness of the final model (i.e., the model norm) in
10
11 210 both horizontal and vertical directions. The synthetic and real models analyzed in this work had their
12
13 211 specific optimal value of λ .
14
15 212

PSO input arguments

16
17 212 The PSO algorithm was iterated enough to guarantee as robust minimization of the fitness
18
19 213 function as possible. Previous PSO applications adopted the maximum number of iterations as the
20
21 214 unique stopping criterion (Godio and Santilano, 2018; Santilano et al., 2018). However, the number
22
23 215 of iterations is problem-dependent and its arbitrary choice can lead to either the ending before the
24
25 216 solution convergence or unnecessary computation (Engelbrecht, 2007). In this work, we took into
26
27 217 account the fitness-function trend during the minimization process. PSO ran as long as the fitness
28
29 218 value did not minimize for 80 consecutive iterations or, if this condition was not satisfied, up to a
30
31 220 maximum number of 6000 iterations. Another stopping criterion was the minimum root mean square
32
33 221 (RMS) error of the data equal to 1 ($\pm 10\%$ of tolerance), in order to avoid the fitting of the data below
34
35 222 their uncertainty (deGroot-Hedlin and Constable, 1990).

36
37 223 The swarm size, i.e., the number of particles, influences the way particles distribute over the
38
39 224 search space to guarantee the exploration of possible solutions. The swarm size must be sufficiently
40
41 225 high to ensure a wide initial coverage of the search space, so that the particles can efficiently explore
42
43 226 all the regions potentially hosting the global minimum. This behavior is missed if the swarm is too
44
45 227 small, giving though the advantage of unburdening the computational complexity. An interesting
46
47 228 analysis on the relation between the swarm size and the computational complexity can be found in
48
49 229 Van den Bergh and Engelbrecht (2001). The number of particles is a problem-dependent parameter
50
51 230 and it is usually set proportional to the number of unknowns, that is for us, the number of resistivity

1
2
3 231 cells the 2D domain was discretized into. The ratio between the problem unknowns and the number
4
5 232 of particles was suggested to be between 8 and 12 times the unknowns by Engelbrecht (2007, p. 241)
6
7 233 for GA and Fernández Martínez et al. (2010a) for PSO. Starting from these guidelines, we performed
8
9 234 a sensitivity analysis to verify the influence of this ratio on the solution of the MT inverse problem.
10
11 235 The number of particles was set 6, 8, 9, 10, and 12 times the number of unknowns. This analysis was
12
13 236 carried out on synthetic model 1 and the results are shown in the next chapter.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The initialization of the optimization is another essential feature of PSO. At the beginning, the particle distribution within the search space is, by default, completely random and bounded between a minimum and maximum value of resistivity. This range is kept constant during the optimization but can vary from each layer (or group of layers or cells) to another (Godio and Santilano, 2018). The decision of the lower and upper resistivity boundaries is problem-dependent and should be coherent with the desired coverage of the search space of solutions. We set the boundaries far larger than the limits of the apparent-resistivity curves. After the random initialization, the adaptive behavior controls the position updating and a stochastic perturbation is guaranteed by γ_1 and γ_2 of equation 1. Local search algorithms usually deploy a starting model (homogeneous or a priori model) to initialize the geophysical inversion. The a priori information is derived from geological (well-log) data or other geophysical methods. Although it is possible to use a priori information to partially influence the swarm behavior, the key factor of global search algorithms like PSO is that they do not require a starting model. To demonstrate this, synthetic data were optimized starting with and without aprioristic information, which was given to the particles in the form of starting positions in the search space. This a priori information was given only to a small amount of particles, 5% of the total, so that the initial position of the rest of the swarm was randomly selected and the swarming nature of PSO was obeyed. We derived the a priori information from the solution PSO gave for the one-dimensional MT inverse problem (Santilano, 2017).

1
2
3 255 The PSO flow chart is shown in Figure 1. This procedure was repeated three times (or “trials”)
4
5 256 for each study case, due to the variability on the solution given by the random initialization. In fact,
6
7 257 the final solutions coming from different initial random distributions are quite similar but not
8 identical, as shown in Santilano et al. (2018) for 1D MT. The solution with the lowest fitness value
9
10 258 was then selected as the final optimized model.
11
12 259
13
14
15 260 **Computational aspect**

17
18 261 Since the optimization process implied the computation of several model responses, the
19
20 262 reliability of the solution was also related to the accuracy of the forward modeling. We adopted the
21
22 263 2D MT forward modeling described in Candansayar (2008) and references therein. It is based on the
23
24 264 finite-difference technique, which solves the complex system of magnetotelluric equations for
25
26 265 Transverse-Electric (TE) and Transverse-Magnetic (TM) polarizations. Firstly, the electric and
27
28 266 magnetic fields are derived for each mesh node and, finally, the apparent resistivity and impedance
29
30 267 phase are calculated. We adopted this forward-modelling code since it is stable, published, and written
31
32 268 in MATLAB.

33
34 269 Addressing the 2D problem made the overall computation time-consuming due to several
35
36 270 factors. The runtime was affected not only by the number of iterations, but also by the population size
37
38 271 and the number of unknowns. The number of iterations depends on the complexity of the problem.
39
40 272 The population size was related to the number of unknowns, i.e., to the desired resolution of the 2D
41
42 273 model. Obviously, a mesh grid unnecessarily dense would have made the computation excessively
43
44 274 long. From taking into account all these issues, a heavy computation effort was to be managed. In
45
46 275 order to speed up the computation, we developed and applied the parallel computing option for the
47
48 276 PSO algorithm. Firstly, we enabled the option “*UseParallel*”, that was potentially provided but not
49
50 277 implemented in the standard code. Then, the most overloaded “*for*” loops were set to run as parallel
51
52 278 *for* loops, such as, for example, the loop that evaluates the fitness function for each particle. Finally,
53
54 279 the PSO algorithm was enabled to run in parallel on the academic cluster by activating the Parallel

1
2
3 280 Computing Toolbox of MATLAB. All the simulations were executed on a 24-core node of a High
4
5 281 Performance Computing (HPC) cluster for academic research. The CPU model of the single node is
6
7 282 an Intel Xeon E5-2680 v3 2.50 GHz (turbo 3.3 GHz) with 128 GB of RAM. The sustained
8
9 283 performance of the cluster is 9.7 TFLOPS.
10
11
12
13 284 2D OPTIMIZATION OF MT SYNTHETIC DATA
14
15
16 285 The theoretical MT data sets were computed from two synthetic models depicted in Figure 2
17
18 286 and Figure 3. They covered a 2D domain 350 km long and 250 km deep, in order to take proper
19
20 boundary conditions into account for the MT forward modeling (Simpson and Bahr, 2005). 15 MT
21
22 stations were centrally placed in the mesh and reciprocally spaced 1.3 km. The mesh discretization
23
24 used for the generation of the synthetic data was different from that used for the optimization.
25
26 Specifically, the latter was slightly coarser than the former due to the computational load given by
27
28 the thousands of forward-modeling calculations in the PSO algorithm. The mesh size along the
29
30 horizontal direction has been kept constant between the stations and doubled from the outer stations
31
32 towards the boundaries. Along the vertical direction, the layer size increased logarithmically with
33
34 depth. In the case of a priori given, the optimization ran on a subdomain of about 400 cells, because
35
36 the a priori information regarded only 15 stations. Without the initial conditioning, the domain
37
38 extended far away from stations due to boundary conditions and the number of cells increased up to
39
40 about 900 for synthetic model 1 and about 750 for synthetic model 2.
41
42
43 298 Both synthetic models simulated the presence of one or more electrically conductive features
44
45 299 embedded in a resistive body beneath the station sites. In case of different configurations, e.g., prior
46
47 knowledge of a conductor outside the location of the sites or the ocean nearby, the mesh discretization
48
49 would have been denser outside the station locations and the mesh boundaries would have been
50
51 enlarged. Moreover, the presence of the ocean could have been addressed in PSO using a priori
52
53 information, as previously explained about the initial position of the particles. The first synthetic
54
55 model is presented in Figure 2. Figure 2a shows the entire 2D mesh, discretized into 33 layers for a
56
57
58 304 model is presented in Figure 2. Figure 2a shows the entire 2D mesh, discretized into 33 layers for a
59
60

1
2
3 305 total of 957 grid cells. The synthetic model 1 is quite simple and composed of a host medium of 100
4 ohm-m including a conductive body of 10 ohm-m from 3 to 5 km depth. Figure 2b shows synthetic
5 306 model 1 as a zoom in the center of the whole mesh.
6
7
8 307
9
10

11 308 Figure 3 illustrates synthetic model 2 as a subsection of the true mesh too. The mesh is
12 discretized into 754 grid cells and, even so, the model space is not under-sampled. The 100 ohm-m
13 309 resistive medium hosts, from the bottom up, a 10 ohm-m body 4-to-9 km deep on the left side of the
14 mesh, a 10 ohm-m body 1-to-2.5 km deep on the right side of the mesh, and a 50 ohm-m body 0-to-
15 310 500 m deep under stations S2-S5. The forward modeling which created the synthetic data considered
16
17 311 26 frequency values between 10^{-2} and 10^3 Hz. Synthetic data were corrupted with uncorrelated
18 Gaussian noise of 10%. This noise corresponded to $\Delta\mathbf{p}_{a,0}$ and $\Delta\Phi_a$ matrices in equation 7, which are
19 the normalization terms of the data misfit. At each k^{th} iteration, the noise influenced the forward
20 calculated response of the corresponding k^{th} model \mathbf{m} . The optimization process was constrained by
21 upper and lower resistivity boundaries equal to 200 ohm-m and 1 ohm-m, respectively.
22
23

24 318 This section is divided into two parts. In the first sub-section, the Lagrange multiplier of
25 synthetic model 1 is identified and, then, the synthetic model 1 is adopted as study case to calibrate
26 two input arguments of PSO, the accelerations and the population size; in the second sub-section, the
27 final resistivity models are presented for both synthetic model 1 and 2.
28
29

322 Calibration of the PSO input arguments

323 The sensitivity analysis on the Lagrange multiplier was carried out on synthetic model 1 using
324 benchmark values for the accelerations and the population size. These values were chosen as benchmarks for
325 the best convergence of the solution after Ratnaweera et al. (2004). The calibration of the accelerations and
326 population size is presented in the next paragraph because it has significance if the most appropriate Lagrange
327 multiplier is adopted. As a benchmark, the cognitive acceleration α_1 linearly decreased from $\alpha_1^{max} = 2$ to α_1^{min}
328 = 0.5 and the social acceleration α_2 linearly increased from $\alpha_2^{min} = 0.5$ to $\alpha_2^{max} = 2$. The benchmark
329 population size was about 9 times the number of unknowns, that is, given 957 cells, 8600 particles. In order to
330
331

1
2
3 330 retrieve the optimal value of the Lagrange multiplier λ , we performed a sensitivity analysis on five different
4 values in the range between 0.001 and 10. λ_x and λ_z were contextually analyzed with the same value and the
5 optimal value was chosen as the point of maximum curvature in the plot of data misfit versus model norm.
6
7 331 Figure 4 shows the data misfit of synthetic model 1 with respect to the model roughness along the horizontal
8 (black diamonds) and vertical (red circles) directions. The best tradeoff value was equal to 0.1 for both λ_x and
9
10 332 λ_z .
11
12 333
13
14 334
15
16

17 335 The sensitivity analysis on the cognitive and social accelerations was carried on for synthetic
18 model 1 once its optimal value of the Lagrange multiplier was identified. For this calibration, the
19 population size was fixed to the aforementioned benchmark value of 8600 particles and its sensitivity
20 analysis is shown in the next paragraph. We chose three different values for the maximum cognitive
21 acceleration, $\alpha_1^{max} = 2.75, 2, 1.5$, and, three different values for the social acceleration, α_2^{min}
22
23 = 0.25, 0.5, 0.75. These values were selected on the basis of the existing literature and equations 5
24 and 6. The solution reliability was evaluated via some parameters of the optimization process, such
25 as the first stopping criterion achieved, the solution clustering, and the trend of the fitness function at
26 each iteration. The simulations ran until one of the three stopping criteria was first fulfilled, that is,
27 when the fitness function did not significantly decrease and (almost) all the particles converged to a
28 unique position in the search space or solutions. Table 1 lists the RMS errors and the fitness-function
29 values ($F(\mathbf{m})$) at the end of the optimization of each test. Our results are largely consistent with the
30 acceleration values pointed out in Ratnaweera et al. (2004) for other applications. The tests using
31 $\alpha_1^k = 1 = 2.75$ with $\alpha_2^k = 1 = 0.5$ and with $\alpha_2^k = 1 = 0.75$ ended before that the RMS was equal to 1,
32 because the fitness function did not decrease for 80 consecutive iterations. These values prevented an
33 effective minimization, as shown in Figure 5, which summarizes the optimization performance using
34 $\alpha_1^k = 1 = 2.75$ and $\alpha_2^k = 1 = 0.5$. The four subplots show, in order: Figure 5a- the fitness-function
35 values of the best particle (red dots) and the mean values of the rest of the swarm (black dots) from
36 the first to the final iteration; Figure 5b- the fitness-function values of the whole swarm as a function
37 of the particle positions in two representative dimensions of the search space, i.e., the first two cells
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 356 of the 2D grid, at the first (grey dots) and final (blue dots) iterations; Figure 5c- the positions, i.e.,
4
5 357 the resistivity values, of the particles in the first two cells of the 2D grid at the first and final iterations
6
7 358 (grey dots and blue dots, respectively); Figure 5d- the histogram containing the distribution of the
8
9 359 fitness-function values at the last iteration among all the particles (8600 in this case). Figure 5 reveals
10
11 360 that the optimization did not end in a convergence state because at the last iteration the minimum
12
13 361 $F(\mathbf{m})$ was not reached by the totality of the particles (Figure 5a and d), and the distribution of the
14
15 362 particles in the search space was still scattered (blue dots in Figure 5b and c). The other tests in Table
16
17 363 1 show an optimal convergence, RMS errors equal to about 1, and the minimized $F(\mathbf{m})$ between 1.33
18
19 364 and 1.73. Figure 6 plots the optimization performance using $\alpha_1^{k=1} = 2$ and $\alpha_2^{k=1} = 0.5$. The
20
21 365 minimization of the fitness function was more effective than that of Figure 5 because all the particles
22
23 366 converged towards a unique position (blue dots in Figure 6b and c) with the same fitness-function
24
25 367 value corresponding to the peak in Figure 6d. It is evident that, iteration by iteration, particles
26
27 368 converged from an initial scattered distribution to a unique position following the best particle
28
29 369 leadership. In this way, the fitness-function value dropped and the histogram developed a unique
30
31 370 peak. This sensitivity analysis confirmed $\alpha_1^{\max} = 2$, $\alpha_1^{\min} = 0.5$, $\alpha_2^{\min} = 0.5$, and $\alpha_2^{\max} = 2$ as optimal
32
33 371 acceleration values for a robust minimization of the fitness function. These accelerations were applied
34
35 372 to the optimization of the other MT data sets of this work.

373 The synthetic model 1 was also a study case for the sensitivity analysis on the population size,
374 in order to assess the influence of the number of particles on both the solution and the runtime. The
375 tests were performed using 5 different values, chosen as multiples of the number of unknowns (957):
376 5700, 7500, 8600, 9500, 11500 particles, that is, 6, 8, 9, 10, and 12 times the unknowns. The
377 accelerations and Lagrange multiplier were set as explained before for the corresponding sensitivity
378 analyses. The results are shown in Table 2. All the tests reached the minimum RMS of about 0.9, but
379 with different numbers of iterations and runtimes because of the different initial distributions of the
380 particles in the search space of solutions. The test using the multiple 8 gave the worst result because

1
2
3 381 the solution was found after the biggest runtime and the highest number of iterations. Differently, the
4
5 382 multiple 9 gave the best result, with the minimum number of iterations and the second shortest
6
7 383 runtime. The ratio of 9, i.e., 8600 particles for synthetic example 1, ensured the most effective
8
9 384 convergence and exploration of the solution space, so that it was adopted for the other tests presented
10
11 385 in this work.
12
13
14
15 386

Results from two synthetic examples

17
18 387 In this sub-section, we present the results of PSO applied to the two synthetic examples
19
20 388 depicted in Figure 2 and Figure 3. The optimization of MT data from the synthetic models was
21
22 389 performed adopting the optimal values for the Lagrange multipliers, accelerations, and population
23
24 390 size reported in the previous sub-section. The results regarding synthetic model 1 are presented in
25
26 391 this order: the model obtained without external conditioning of the PSO initialization, the model
27
28 392 resulting from a poorly populated swarm, and finally the result after the PSO initialization with a
29
30 393 priori information given as starting model. The resistivity model obtained without a priori information
31
32 394 is shown in Figure 7. After about 150 iterations and 4.6 hours, the RMS error stabilized around the
33
34 395 final value of 0.86, while $F(\mathbf{m})$ was 1.37. These values are listed in Table 3, but they can also be
35
36 396 found in Table 1 for $\alpha_1^{k=1} = 2$ and $\alpha_2^{k=1} = 0.5$, and in Table 2 for 8600 particles. The resistivity
37
38 397 model in Figure 7 was largely comparable to the true model in Figure 2, since the conductive anomaly
39
40 398 was correctly detected in both size and resistivity. Figure 8 plots the fitting curves between synthetic
41
42 399 and calculated data for both apparent resistivity (ρ_{app}) and phase, both TE and TM polarizations. The
43
44 400 synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are
45
46 401 plotted as solid lines for TE, and dashed lines for TM. Four stations were selected for their different
47
48 402 positions in relation to the lateral discontinuities: S1, S4, S7, and S11. They show an example of poor
49
50 403 ($S4 \rho_{app}$), average ($S11$), and good fit ($S1$, and $S7$). Considering the high number of unknowns and
51
52 404 the wide range of variation of ρ_{app} , it could be said that these curves are clearly similar to each other
53
54 405 as also proved by the low RMS error.

1
2
3 406 The influence of the population size on the optimization process is presented in Table 2. Figure
4
5 407 9 shows the effect of a poorly populated swarm on the final resistivity model. This result followed
6
7 408 from a population size of 5700 particles, i.e., 6 times the unknowns. The result was similar to the true
8
9 409 model in Figure 2, since the conductive body was identified. However, the output was not completely
10
11 410 appreciable due to some lateral conductive artefacts that broke in the homogeneous 100-ohm-m
12
13 411 background. As expected, this outcome was the consequence of an ineffective initial distribution of
14
15 412 the particles in the search space of solutions, and, possibly, of the missing of the global minimum.

16
17
18 413 The a priori information used to initialize the optimization came from the PSO solutions of
19
20 414 the 1D inverse problem for the 15 stations of synthetic model 1. Only 5% of the particles were initially
21
22 415 influenced with this solution. After 250 iterations, the RMS error reached the minimum threshold,
23
24 416 with a corresponding fitness-function value of 1.4. The final resistivity model is shown in Figure 10
25
26 417 and is comparable to the original of Figure 2, since the conductive anomaly was adequately identified.
27
28 418 Figure 11 plots the fitting curves of the selected stations. The PSO-predicted responses were distinctly
29
30 419 consistent with the synthetic data and the difference with the curves of Figure 8 was negligible, except
31
32 420 for the slight improvement for ρ_{app} of S4 and S11. Table 3 lists the details regarding RMS, runtime (in
33
34 421 hours), and the total number of iterations.

35
36 422 The optimization of synthetic data from synthetic model 2 (true model in Figure 3) was
37
38 423 performed after the calibration of the input arguments, the accelerations, and population size. The
39
40 424 identification of the optimal Lagrange multiplier λ for synthetic model 2 was inferred from the L-curve
41
42 425 response presented in Figure 12. It refers to the data-misfit trend as a function of the horizontal (black
43
44 426 diamonds) and vertical (red circles) roughness for synthetic model 2. The best tradeoff value was 0.1
45
46 427 for both λ_x and λ_z .

47
48 428 The best solution for synthetic model 2 without a priori initialization is illustrated in Figure
49
50 429 13. All the three low-resistivity bodies were accurately positioned as can be seen from the zoom-in
51
52 430 panel. After 1674 iterations, the minimum fitness function value was 1.3 and the RMS error was 0.9.

1
2
3 431 Figure 14 plots the comparison between synthetic and calculated data for both apparent resistivity
4
5 432 (ρ_{app}) and phase (TE and TM). Stations S1, S5, S9, and S12 were selected as representative for their
6
7 433 poor (S1), average (S5 and S9), and good fit (S12). Taking the complexity of this synthetic example
8
9 434 into proper account, the curves are in good agreement. The runtime is reported in Table 3.
10
11
12

13 435 The a priori information was set as previously explained. Once 5% of the particles were
14
15 436 initially influenced, convergence was reached only after 53 iterations with a minimum fitness-
16
17 437 function value of 1.5. Figure 15 displays the final output. The 10-ohm-m lateral bodies were correctly
18
19 438 imaged, while the superficial 50-ohm-m body was scarcely identified. A distinct difference between
20
21 439 the models with and without a priori is indeed the superficial body, as can be seen in the zoom-in
22
23 440 panels of Figure 13 and Figure 15. Figure 16 graphically demonstrates the low RMS error of 0.99,
24
25
26 441 which is reported in Table 3. The fitting curves of Figure 16 show a good agreement between synthetic
27
28 442 and predicted data and no significant improvements compared to Figure 14.

443 2D OPTIMIZATION OF MT FIELD DATA

444 The COPROD2 data set

445 36 445 The COPROD2 data set collects long-period MT measurements along a profile of 35 sites
37
38 446 crossing a 2D geoelectrical structure in Saskatchewan and Manitoba, Canada (Jones and Savage,
39
40 447 1986). The name stands for “Comparison of One-dimensional PROfiles from MT Data”, while the
41
42 448 “2” refers to the two-dimensionality, differently from the one-dimensional data set called
43
44 449 “COPROD”. The most appreciable advantages of this data set are the following: A wide period
45
46 450 bandwidth (from $2.6 \cdot 10^{-3}$ s to $1.8 \cdot 10^3$ s), low impedance errors (< 2%), previous static shift correction,
47
48 451 and the possibility of comparing different models from well-established inversion algorithms (Jones,
49
50 452 1993b). Our aim is to apply the PSO algorithm to detect deep electromagnetic anomalies, while any
51
52 453 geological interpretation is beyond the scope of this paper.

1
2 Since responses at low periods (below 10 s) have been widely recognized as one-dimensional,
3

4 454 original data were selected from 10.67 s to 910.2 s (deGroot-Hedlin and Constable, 1993). As
5

6 455 proposed in the aforementioned studies, a subset of 20 MT stations, from the 8th to the 27th of the
7

8 456 original line, was chosen to focus only on the center of the 400 km east-west profile. This selection
9

10 457 was adopted also because these 20 sites have the same number of acquisition frequencies (14) within
11

12 458 the considered interval. The errors on the data were kept as original for both TE and TM apparent
13

14 459 resistivity and phase. Figure 17 plots MT observations and error bars for stations 12, 13, and 14,
15

16 460 chosen as representative of the subset. The maximum observed error is 1.2 for TE apparent resistivity
17

18 461 (on logarithmic scale) and 6.92° for TM phase.
19

20 462 The 2D model was divided into 10 layers, from 1.8 km to 60.5 km deep, and the thickness of
21

22 463 each layer increased logarithmically with depth. Along the horizontal direction, the mesh was about
23

24 464 200 km long and subdivided into 34 bricks, one for each station plus others as boundary conditions.
25

26 465 The total number of cells was 340. Since some structures of the region are known to be highly
27

28 466 conductive, the lower boundary of the problem was set equal to 0.1 ohm-m. Literature references also
29

30 467 state that superficial sediments are far more conductive than the resistive basement. For this reason,
31

32 468 the upper boundary of resistivity was chosen to be different between the upper and underlying layers.
33

34 469 We observed that a search space too large for the upper layers would have driven the solution toward
35

36 470 no convergence and erroneous local minima. In detail, the first two superficial layers, namely up to 5
37

38 471 km deep, had 10 ohm-m as upper boundary, while, the layers below, 1000 ohm-m. The population
39

40 472 size was equal to 2500 particles, proportional to the number of cells. The Lagrange multiplier λ was
41

42 473 chosen after a sensitivity analysis on five different values in the range between 0.001 and 10. The
43

44 474 value that coincides with the point of maximum curvature in the plot of data misfit versus model norm
45

46 475 is 0.1, as shown in Figure 18. A priori information was not given: The optimization started with a
47

48 476 completely random initialization.
49

50 477

51 478

52 479

53 480

54 481

55 482

56 483

57 484

58 485

59 486

60 487

1
2
3 478 The final model from COPROD2 data was computed after 6000 iterations and is depicted in
4
5 479 Figure 19. The shallow conductive structure was extensively identified, while, at depth from 5 km to
6
7 480 the bottom, the background resistivity was predominantly 1000 ohm-m. The most significant feature
8
9 481 of this model was represented by the low-resistivity anomalies below the station E3-E4 and 12-11 at
10
11 482 around 20 to 35 km of depth. Our output is well comparable with the ones represented in Figure 20
12
13 483 and reported by Jones (1993b). For ease of comparison, both the color scale and the name of the
14
15 484 stations of Figure 19 were plotted as the original ones in Figure 20. There was good agreement
16
17 485 between our model and those called “de groot-2” (deGroot-Hedlin and Constable, 1993), “rasmussen”
18
19 486 (Rasmussen, 1993), “wu” (Wu et al., 1993), and “uchida” (Uchida, 1993): Low-resistivity anomalies
20
21 487 were identified in the same regions. The similarity can be explained by the same approach adopted in
22
23 488 the Occam’s inversion using the smoothing parameter, excepted for “wu”, which used a different
24
25 489 approach. In detail, the most evident similarity was the conductive region in the first 5-7 km of depth.
26
27 490 Another similarity regarded the 30-km-deep conductor below the stations from E2 to 14 and its
28
29 491 extension at greater depths. A further correspondence was the interruption of the 1000 ohm-m
30
31 492 structure below the stations from 13 to 11 at about 20 km of depth. Apart from the “wu” model, all
32
33 493 the analyzed results presented a low-resistivity region (about 100 ohm-m) in the westernmost part of
34
35 494 the model at a depth greater than 30 km.

495 Figure 21 plots the apparent resistivity (ρ_{app}) and phase at selected periods for the 20 stations
496 in the horizontal axis. The observed data are marked with dots for TE and diamonds for TM, and
497 predicted responses are plotted with a solid line for TE and a dashed line for TM. The RMS error is
498 reported in Table 3. The final fitness-function value was 26.6.

53 499 Computation time

54
55 500 The computationally demanding nature of the PSO algorithm was actually expected due to the high
56
57 number of: i) iterations, ii) population size, and iii) cells assembling the mesh. The standard release
58 501 of the code was not effective in addressing the 2D inverse problem, therefore we applied some
59
60 502

1
2
3 503 modifications to develop a parallelized version of PSO. The tests on a HPC cluster proved that, when
4
5 504 24 cores were adopted, the runtime saving was more than 80% with respect to the use of 4 cores (see
6
7 505 Figure 22). A test using the non-parallelized release of the code (one single worker) would have been
8
9 506 unfeasible in terms of machine working load. Figure 22 shows the dramatic speedup of PSO
10
11 507 computation for a reference simulation of 150 iterations and 10000 particles: Black lines indicate the
12
13 508 running duration in hours, while blue lines the total speedup in percentage. The parallel environment
14
15 509 “shared” (dotted lines) exploited workers of the same node, while “orte” (dashed lines) referred to
16
17 510 workers from different machines of the cluster. It could be seen that “shared” was a bit faster than
18
19 511 “orte”, especially at high numbers of cores.

20
21
22 512 The total runtime of PSO computations is reported in Table 3 for each study case. These values
23
24 513 refer to one single trial, while a total of three trials were performed. Runtimes are not directly
25
26 514 comparable to each other because the stopping criterion was met after different numbers of iterations,
27
28 515 that is, less than 1600 iterations for the synthetic models and 6000 for the real data set. The
29
30 516 optimization of synthetic examples stopped because the minimum RMS error was achieved, while
31
32 517 the optimization of the real data stopped because the fitness function did not minimize for 80
33
34 518 iterations. The synthetic examples were optimized in a few iterations, but the runtime was longer than
35
36 519 that of COPROD2 data set due to the higher number of layers and particles. In fact, the big difference
37
38 520 between the computation times is explained by the low number of layers of the COPROD2 model is
39
40 521 discretized -about one-third of that of the synthetic models. The optimization of synthetic data 1 and
41
42 522 2 without a priori differed in the number of iterations needed, but the runtimes are quite similar, if the
43
44 523 proportion between the iterations is taken into account. The optimization of synthetic data 1 and 2
45
46 524 had shorter runtimes when a priori initialization was given.

525 DISCUSSION

526 Our tests on synthetic data demonstrate the reliability of PSO in solving the 2D inverse
527 problem for MT data sets.

1
2
3 528 The choice of the most appropriate values of accelerations and population size was crucial for
4
5 529 obtaining valid models. The initial sensitivity analysis on the PSO input arguments was essential to
6
7 530 identify the most appropriate tuning coefficients which effectively minimized the objective function
8
9 531 and enhanced the solution convergence. The calibration of the social and cognitive accelerations led
10
11 532 to the optimal values of $\alpha_1^{\max} = 2$, $\alpha_1^{\min} = 0.5$, $\alpha_2^{\min} = 0.5$, and $\alpha_2^{\max} = 2$. Our findings are hence in
12
13 533 agreement with Ratnaweera et al. (2004). We demonstrated that the population size was directly
14
15 534 proportional to the total runtime, but, at the same time, a poorly populated swarm negatively
16
17 535 influenced the model. The reason for this was that the search space was ineffectively covered by the
18
19 536 initial random distribution of the particles. We showed that the best ratio between the number of
20
21 537 unknowns and the number of particles was 9. This outcome is significant for our high-dimensional
22
23 538 problem because, so far, the literature has suggested increasing the number of particles up to 12 times
24
25 539 the number of unknowns. The conclusion of our analysis slightly modifies this ratio, with the
26
27 540 advantage of avoiding extra computational load.

541 We showed that the application of PSO did not require an initial assumption about the solution
542 (i.e., a priori information). At the same time, we introduced a novel and valid tool to potentially
543 communicate external or additional information to the swarm, in terms of the initial position of
544 particles within the search space. Our findings showed that, if the geological or geophysical
545 information is reliable (e.g., from wells, seismic reflectors and so on), it influences the behavior of
546 particles at the beginning of the optimization. This kind of initialization resulted in shorter runtimes
547 because the swarm did not waste time searching for local minima, which were already given from the
548 beginning (see Table 3). On the other hand, using default random initialization, the results of synthetic
549 models proved that there was no requirement for a priori initialization, since the final resistivity model
550 was perfectly comparable with the original synthetic model. Moreover, there was high solution
551 quality, despite two factors: Gaussian noise disturbing the data and the presence of equivalent
552 solutions in the MT inverse problem. In fact, the conductive anomalies embedded in the resistive host
553
554
555
556
557
558
559
560

1
2
3 553 medium were accurately identified in terms of size and resistivity values. RMS errors were around 1
4
5 554 and, interestingly quite similar with and without a priori initialization. Another element confirming
6
7 555 the PSO independence from the starting model is the comparison of the fitting curves in Figure 8 and
8
9 556 Figure 11 for synthetic model 1, and Figure 14 and Figure 16 for synthetic model 2. These plots
10
11 557 proved that there were not substantial differences between calculated responses with and without a
12
13 558 priori, even considering the stations above the lateral discontinuities (S4 and S11 for synthetic model
14
15 559 1 and S5 and S9 for synthetic model 2).
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

20 560 A significant result arose from the application of PSO to real 2D data, the COPROD2 data
21
22 set. As regards the problem settings, the uppermost layers of the 2D mesh had different boundary
23
24 conditions with respect to the underlying layers, due to the complexity of the investigated area. Many
25
26 applications of global search algorithms to geophysics have considered different resistivity
27
28 boundaries between one layer (or group of layers) and another, so that each unknown of the problem
29
30 can independently be bounded within its search space (Godio and Santilano, 2018). The setting of the
31
32 boundary conditions is not a trivial step for the deterministic inversion either, because it implies full
33
34 comprehension of the problem and some insight into the possible solution. For the optimization of
35
36 the COPROD2 data set, preliminary information from the geology of the area facilitated the definition
37
38 of a wide interval of resistivity values within which the solution could be searched. We applied this
39
40 approach in order to enhance the convergence of the solution. This option distinguished the solution
41
42 space of upper layers from that of deep layers. The final model had a mean resistivity of 6 ohm-m in
43
44 the two upper layers, thus confirming the presence of superficial sediments. At a depth of about 25
45
46 km, a conductive region breaks the 1000 ohm-m background, with a minimum value of 1.2 ohm-m.
47
48
49 573 The final RMS error was slightly bigger than that of the synthetic examples and negatively affected
50
51 by the mismatch of data at long periods (Figure 21). This was unexpected, given the satisfactory
52
53 behavior of the fitting curves of the synthetic examples in Figure 8 and Figure 14. It would have
54
55 been interesting to quantitatively compare our result with the model obtained by Everett and Schultz
56
57
58
59
60

1
2
3 578 (1993) using GA, which is a global search algorithm too. Unfortunately, their RMS Misfit of 1.48 is
4
5 579 not directly comparable with our value of 2.42 because there were substantial differences in the
6 method, such as: Period range, number of stations, mesh discretization, stopping criteria for iterations.
7
8 580 Interestingly, the adoption of Occam-like optimization may provide a more effective solution of the
9 resistivity distribution with respect to the GA. It has also been proved in literature that PSO ensures
10 581 a higher convergence with respect to the other global search algorithms (Yuan et al., 2009; Fernández
11
12 582 Martínez et al., 2010a). The application to field data represents a new encouraging approach for their
13
14 583 optimization by means of computational swarm intelligence.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

As regards the optimizations run without a priori initialization, the runtimes of the different data sets were not straightforwardly comparable due to the different number of iterations required to achieve convergence. The COPROD2 optimization needed 6000 iterations to stop at RMS=2.42, while the synthetic examples reached RMS=1 in fewer iterations but taking a runtime proportionally longer than that for the real-data optimization. This is mainly explained by the high level of mesh discretization for the synthetic models, about 800 cells, compared to about 340 cells for the domain chosen for real-data interpretation. The more the unknowns, the greater the swarm size and hence the computation time. The computation speedup was obtained by introducing the parallelization of the code. The tests performed on a HPC cluster pointed out the capability of our version of PSO to speed up the computation by more than 4 times with respect to running it on a simple machine of 4 cores. The decrease of the runtime allowed us to efficiently perform several trials of the optimization process, starting from different random distributions of the swarm. The final fitness-function value of the synthetic examples was lower than that of the real data due to the peculiarity of the data sets. The choice of the optimal Lagrangian multiplier may be seen as a computational cost, because the sensitivity analysis of different values of λ was performed. This analysis could represent a slight limitation of the presented method, since PSO ran for each investigated value of λ . However, once the balance was found, we were able to deploy the model with the adequate level of smoothing.

Although we reduced the computation time, it remained not comparable with that of deterministic algorithms. PSO applied to the 2D inverse problem is relatively time-consuming if parallelization cannot be exploited and densely discretized meshes are adopted. However, the parallelization code has the potential of making the PSO computation more manageable. Moreover, high computing capacity is nowadays within everyone's reach and global search algorithms, despite the critical view of the past, can now be considered worthy of attention. We do not see the long runtime as a scientific barrier for the application of PSO to high-dimensional geophysical problems. The computational load was balanced by the advantages of this meta-heuristic method, namely, the independence from the choice of the starting model and the solution driven by the evolutionary search.

CONCLUSION

The Particle Swarm Optimization (PSO) algorithm has proven to be a valid method to solve the two-dimensional (2D) inverse problem for magnetotelluric (MT) data, for both synthetic and field (COPROD2) data sets. This work extended the application of PSO to MT inversion from the one-dimensional problem, already visited in the literature, to the 2D problem. The stochastic nature of PSO and the combination of exploration and exploitation behaviors played a key role in finding the optimized solution within the search space, which was composed of all the possible solutions of resistivity models.

The standard release of the code was easily implemented for our specific application. We observed striking improvements moving from standard PSO to hierarchical PSO with time-varying acceleration coefficients (HPSO-TVAC). This issue has not been addressed in previous research on PSO applied to geophysics, but was crucial in the optimization of 2D MT data. In fact, thanks to time-varying acceleration coefficients, the optimization ended with true convergence and stability. The complexity of the 2D problem had a direct influence on the computation time, which we reduced with

1
2
3 627 the parallelization of the code. Running PSO on a High Performance Computing (HPC) cluster
4
5 628 resulted in runtime savings of about 80%.
6
7
8 629
9
10 630
11
12 631
13
14 632
15
16 633
17
18 634
19
20 635
21
22 636
23
24 637
25
26 638
27
28 639
29
30 640

We carried out a detailed sensitivity analysis on some input parameters of the PSO algorithm due to their direct influence on the stability and convergence of the solution. The social and cognitive accelerations and the population size were investigated to retrieve their optimal values and analyze their effect on both the final resistivity models and total runtime. We first applied PSO to 2D MT synthetic data, in order to validate the method. The initialization of the optimization was purely random by default, but we also tried to influence it with a starting model derived from PSO solutions of the 1D problem. In this case, the optimization was externally but not totally influenced, because only a small portion of the swarm was initialized. We proved that a priori information as the starting model can be avoided. The resistivity models which did not receive the a priori initialization were in line with the original synthetic models. Then, PSO was applied to the field data set COPROD2, with a random initialization. The optimization of COPROD2 data produced a valid resistivity model, largely comparable with results from existing research.

31
32 641 The most important conclusions of this work are that PSO can be successfully applied to the
33
34 642 2D MT inverse problem and the a priori starting model is not required for the achievement of valid
35
36 643 models. Our results are encouraging enough to extend the application of evolutionary algorithms to
37
38 644 other geophysical inverse problems, bearing in mind that the high dimensionality of the problem
39
40 645 implies runtimes longer than those of local search methods. Future work will investigate a further
41
42 646 speedup of the PSO computation, as well as its application to other MT real data sets. A possible
43
44 647 direction of future studies should consider the comparison between local and global search methods
45
46 648 for the 2D MT inversion, as well as the challenge of the 3D inverse problem.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

AKNOWLEDGEMENTS

The authors thank the editors and the four reviewers for their valuable and helpful comments, which resulted in a vastly improved paper. Computational resources provided by hpc@polito (<http://hpc.polito.it>).

REFERENCES

- Avdeev, D.B., 2005, Three-Dimensional electromagnetic modelling and inversion from theory to application: *Surveys in Geophysics*, **26**, 6, 767-799. doi: 10.1007/s10712-005-1836-2.
- Boerner, R. U., 2010, Numerical modelling in geo-electromagnetics: Advances and challenges: *Surveys in Geophysics*, **31**, 2, 225-245. doi: 10.1007/s10712-009-9087-x.
- Brodie, R. C., and W. Jiang, 2018, Trans-dimensional Monte Carlo inversion of short period Magnetotelluric data for cover thickness estimation: ASEG Extended Abstracts, 1-7. doi: 10.1071/ASEG2018abT5_1F
- Candansayar, M. E., 2008, Two-dimensional inversion of magnetotelluric data with consecutive use of conjugate gradient and least-squares solution with singular value decomposition algorithms: *Geophysical Prospecting*, **56**, 1, 141-157.
- Conway, D., J. Simpson, Y. Didana, J. Rugari, and G. Heinson, 2018, Probabilistic Magnetotelluric inversion with adaptive regularisation using the No-U-Turns sampler: *Pure and Applied Geophysics*, **175**, 8, 2881-2894. doi: <https://doi.org/10.1007/s00024-018-1870-5>
- Darisma, D., U. Said, and W. Srigutomo, 2017, 2D Gravity Inversion Using Particle Swarm Optimization Method: 23rd European Meeting of Environmental and Engineering Geophysics, EAGE, Extended Abstracts, We 23P2 16.
- deGroot-Hedlin, C., and S. Constable, 1990, Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data: *Geophysics*, **55**, 12, 1613-1624.
- deGroot-Hedlin, C., and S. Constable, 1993, Occam's inversion and the North American Central Plains electrical anomaly: *Journal of geomagnetism and geoelectricity*, **45**, 9, 985-999.
- Dong, H., and A. G. Jones, 2018, On the influences of random starting/prior models in Three-Dimensional Magnetotelluric inversion: 24th EM Induction Workshop, abstract.
- Dosso, S., and D. Oldenburg, 1991, Magnetotelluric appraisal using simulated annealing: *Geophysical Journal International*, **106**, 2, 379-385.
- Ebbesen, S., P. Kiwitz, and L. Guzzella, 2012, A generic particle swarm optimization matlab function: *American Control Conference, IEEE, Extended Abstracts*, 1519-1524.
- Engelbrecht, A. P., 2007, Computational Intelligence An Introduction: John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, England.
- Everett, M., and A. Schultz, 1993, Two-dimensional nonlinear magnetotelluric inversion using a genetic algorithm: *Journal of geomagnetism and geoelectricity*, **45**, 9, 1013-1026.
- Farquharson, C. G., and D. W. Oldenburg, 2004, A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems: *Geophysical Journal International*, **156**, 3, 411-425.

- 1
2
3 687 Fernández Martínez, J. L., E. García Gonzalo, J. P. Fernández Álvarez, H. A. Kuzma, and C.
4 688 O. Menéndez Pérez, 2010a, PSO: A powerful algorithm to solve geophysical inverse problems:
5 689 Journal of Applied Geophysics, **71**, 1, 13-25. doi: 10.1016/j.jappgeo.2010.02.001.
6 690 Fernández Martínez, J. L., E. García-Gonzalo, and V. Naudet, 2010b, Particle swarm
7 691 optimization applied to solving and appraising the streaming-potential inverse problem: Geophysics,
8 692 **75**, 4, WA3-WA15, doi: 10.1190/1.3460842.
9
10 693 Ghaedrahmati, R., A. Moradzadeh, N. Fathianpour, and S. K. Lee, 2014, Investigating 2-D
11 694 MT inversion codes using real field data: Arabian Journal of Geosciences, **7**, 6, 2315-2328.
12
13 695 Godio, A., and A. Santilano, 2018, On the optimization of electromagnetic geophysical data:
14 696 Application of the PSO algorithm: Journal of Applied Geophysics, **148**, 163-174.
15 697 Jones, A. G., 1993a, COPROD2 Data set accessed October 2017 at <http://www.complete-mt-solutions.com/mtnet/data/coprod2/coprod2.html>.
16
17 698 Jones, A. G., 1993b, The COPROD2 dataset: Tectonic setting, recorded MT data, and
18 699 comparison of models: Journal of geomagnetism and geoelectricity, **45**, 9, 933-955.
19
20 700 Jones, A. G., and P. J. Savage, 1986, North American Central Plains conductivity anomaly
21 701 goes east: Geophysical Research Letters, **13**, 7, 685-688.
22
23 702 Kennedy, J., and R. Eberhart, 1995, Particle Swarm Optimization: International Conference
24 703 on Neural Networks IV, IEEE, Extended Abstracts, 1000, 1942-1948.
25
26 704 Kennedy, J., R. Eberhart, and Y. H. Shi, 2001, Swarm intelligence: Morgan Kaufmann
27 705 Publishers, San Mateo, CA.
28
29 706 Miensopust, M. P., P. Queralt, A. G. Jones, and 3D MT modellers, 2013, Magnetotelluric 3-
30 707 D inversion—a review of two successful workshops on forward and inversion code testing and
31 708 comparison: Geophysical Journal International, **193**, 3, 1216-1238.
32
33 709 Newman, G., 2014, A review of high-performance computational strategies for modeling and
34 710 imaging of electromagnetic induction data: Surveys in Geophysics, **35**, 1, 85-100. doi:
35 711 10.1007/s10712-013-9260-0.
36
37 712 Olalekan, F., and Q. Di, 2017, Particle Swarm Optimization method for 1D and 2D MTEM
38 713 data inversion: SEG International Exposition and 87th Annual Meeting, SEG, Extended Abstracts,
39 714 1219-1224.
40
41 715 Pace, F., A. Santilano, and A. Godio, 2017, Particle Swarm Optimization of Electromagnetic
42 716 Data with Parallel Computing in the 2D Case: 23rd European Meeting of Environmental and
43 717 Engineering Geophysics, EAGE, Extended Abstracts, Tu 23 B04.
44
45 718 Perez, R. E., and K. Behdinan, 2007, Particle swarm approach for structural design
46 719 optimization: Computers and Structures, **85**, 1579-1588.
47
48 720 Pérez-Flores, M. A., and A. Schultz, 2002, Application of 2-D inversion with genetic
49 721 algorithms to magnetotelluric data from geothermal areas: Earth, planets and space, **54**, 5, 607-616.
50
51 722 Rasmussen, T. M., 1993, Two-dimensional Occam model of COPROD2 data-first order
52 723 description of resolution and variance: Journal of geomagnetism and geoelectricity, **45**, 9, 1027-1037.
53
54 724 Ratnaweera, A., S. K. Halgamuge, and H. C. Watson, 2004, Self-organizing hierarchical
55 725 particle swarm optimizer with time-varying acceleration coefficients: IEEE Transactions on
56 726 evolutionary computation, **8**, 3, 240-255.
57
58 727 Santilano, A. 2017, Deep geothermal exploration by means of electromagnetic methods: New
59 728 insights from the Larderello geothermal field (Italy): PhD thesis, Politecnico di Torino.
60

- 1
2
3 730 Santilano, A., A. Godio, and A. Manzella, 2018, Particle swarm optimization for simultaneous
4 731 analysis of Magnetotelluric (MT) and Time Domain EM (TDEM) data: *Geophysics*, **83**, 3, E151-
5 732 E159.
6
7 733 Sen, M. K., and P. L. Stoffa, 2013, Global optimization methods in geophysical inversion:
8 734 Cambridge University Press.
9
10 735 Shaw, R., and S. Srivastava, 2007, Particle swarm optimization: A new tool to invert
11 736 geophysical data: *Geophysics*, **72**, 2, F75-F83.
12
13 737 Shi, Y., and R. Eberhart, 1998, A modified particle swarm optimizer: The 1998 IEEE
14 738 International Conference on Evolutionary Computation IEEE, Extended Abstracts, 69-73.
15
16 739 Simpson, F., and K. Bahr, 2005, Practical magnetotellurics: Cambridge University Press.
17
18 740 Siripunvaraporn, W., 2012, Three-dimensional magnetotelluric inversion: An introductory
19 741 guide for developers and users: *Surveys in Geophysics*, **33**, 1, 5-27. doi: 10.1007/s10712-011-9122-
20 742 6.
21
22 743 Uchida, T., 1993, Inversion of COPROD2 magnetotelluric data by use of ABIC minimization
23 744 method: *Journal of geomagnetism and geoelectricity*, **45**, 9, 1063-1071.
24
25 745 Van den Bergh, F., and A. P. Engelbrecht, 2001, Effect of swarm size on cooperative particle
26 746 swarm optimizers: Genetic Evolutionary Computation Conference (GECCO-2001), Proceedings,
27 747 892-899.
28
29 748 Wu, N., J. R. Booker, and J. T. Smith, 1993, Rapid two-dimensional inversion of COPROD2
30 749 data: *Journal of geomagnetism and geoelectricity*, **45**, 9, 1073-1087.
31
32 750 Xiang, E., R. Guo, S. E. Dosso, J. Liu, H. Dong, and Z. Ren, 2018, Efficient hierarchical trans-
33 751 dimensional Bayesian inversion of magnetotelluric data: *Geophysical Journal International*, **213**, 3,
34 752 1751-1767. doi: <https://doi.org/10.1093/gji/ggy071>
35
36 753 Yan, P., M. A. Garcia Juanatey, T. Kalscheuer, C. Juhlin, P. Hedin, A. Savvaidis, H. Lorenz,
37 754 and J. Kück, 2017a, A magnetotelluric investigation of the Scandinavian Caledonides in western
38 755 Jämtland, Sweden, using the COSC borehole logs as prior information: *Geophysical Journal
39 756 International*, **208**, 3, 1465-1489. doi: <https://doi.org/10.1093/gji/ggw457>.
40
41 757 Yan, P., T. Kalscheuer, P. Hedin, and M. A. Garcia Juanatey, 2017b, Two-dimensional
42 758 magnetotelluric inversion using reflection seismic data as constraints and application in the COSC
43 759 project: *Geophysical Research Letters*, **44**, 3554-3563. doi: 10.1002/2017GL072953.
44
45 760 Yuan, S., S. Wang, and N. Tian, 2009, Swarm intelligence optimization and its application in
46 761 geophysical data inversion: *Applied Geophysics*, **6**, 2, 166-174. doi: 10.1007/s11770-009-0018-x.
47
48 762 Zhan, Z.-H., J. Zhang, Y. Li, and H. S.-H. Chung, 2009, Adaptive particle swarm
49 763 optimization: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **39**, 6,
50 764 1362-1381.
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5

CAPTIONS

6
7

Figure 1. The PSO algorithm flowchart: \mathbf{P} is the local best solution; \mathbf{G} is the global best solution.

8
9
10

Figure 2. Synthetic model 1: a) the 2D mesh is discretized into 33 layers and a total of 957 grid cells.

11
12
13
14
15

The labels $S1, \dots, S15$ indicate the location of the 15 MT stations. The dashed area is shown in b): A 10 ohm-m conductive body is hosted in a 100 ohm-m medium.

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 3. Synthetic model 2: Two 10 ohm-m deep anomalies and one superficial 50 ohm-m body are embedded in a 100 ohm-m host medium. The labels $S1, S2, \dots, S15$ indicate the 15 MT stations. The zoomed-in box on the top shows the 50-ohm-m body below $S2-S5$.

Figure 4. L-curve response for synthetic model 1 along the horizontal (black diamonds) and vertical (red circles) directions. The tradeoff between data misfit and model norm indicates the best Lagrange multiplier λ equal to 0.1.

Figure 5. Fitness function $F(\mathbf{m})$ and particle positions at the end of the optimization: a) fitness-function value, iteration after iteration, for the best particle (red dots) and the rest of the swarm (black dots); b) the fitness-function value as a function of the particle positions in the resistivity (ρ) search space, at the first (grey dots) and final (blue dots) iterations; c) plain view of b); d) final distribution of the fitness-function values among all the particles.

Figure 6. Fitness function $F(\mathbf{m})$ and particle positions at the end of the optimization: a) fitness-function value, iteration after iteration, for the best particle (red dots) and the rest of the swarm (black dots); b) the fitness-function value as a function of the particle positions in the resistivity (ρ) search space, at the first (grey dots) and final (blue dot) iterations; c) plain view of b) with all particles converged to the last position (red circled blue dot); d) final distribution of the fitness-function values among all the particles.

Figure 7. PSO solution for synthetic model 1, after about 150 iterations without a priori initialization for the 8600 particles of the swarm. Lagrange multiplier $\lambda = 0.1$.

Figure 8. Fitting curves between data of synthetic model 1 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are $S1, S4, S7$, and

1
2
3 26 *S11. The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are*
4
5 27 *plotted as solid lines for TE, and dashed lines for TM. The optimization was randomly initialized.*
6
7

8 28 *Figure 9. PSO solution for synthetic model 1 using a swarm size of only 5700 particles (6 times the*

9
10 29 *unknowns), after about 160 iterations, without a priori initialization, and Lagrange multiplier $\lambda = 0.1$.*
11
12

13 30 *Figure 10. PSO solution for synthetic model 1, after about 250 iterations and with a priori information*

14 31 *given to 5% of the particles. Lagrange multiplier $\lambda = 0.1$.*
15
16

17 32 *Figure 11. Fitting curves between data of synthetic model 1 and calculated data for apparent resistivity*

18 33 *(ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are S1, S4, S7, and*

19 34 *S11. The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are*

20 35 *plotted as solid lines for TE, and dashed lines for TM. The optimization was initialized with a priori*

21 36 *information.*
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

37 37 *Figure 12. L-curve response for synthetic model 2 along horizontal (black diamonds) and vertical (red*

38 38 *circles) directions. The tradeoff between data misfit and model norm indicates the best Lagrange multiplier λ*

39 39 *equal to 0.1.*
40
41

42 42 *Figure 13. PSO solution for synthetic model 2, after 1674 iterations and without a priori initialization.*

43 43 *Lagrange multiplier $\lambda=0.1$.*
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

42 42 *Figure 14. Fitting curves between data of synthetic model 2 and calculated data for apparent*

43 43 *resistivity (ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are S1, S5,*

44 44 *S9, and S12. The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted*

45 45 *data are plotted as solid lines for TE, and dashed lines for TM. The optimization was randomly initialized.*
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

46 46 *Figure 15. PSO solution for synthetic model 2, after 53 iterations and with a priori information given*

47 47 *to 5% of the particles. Lagrange multiplier $\lambda = 0.1$.*
48
49
50
51
52
53
54
55
56
57
58
59
60

48 48 *Figure 16. Fitting curves between data of synthetic model 2 and calculated data for apparent resistivity*

49 49 *(ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are S1, S5, S9, and*

50 50 *S12. The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are*

1
2
3 51 plotted as solid lines for TE, and dashed lines for TM. The optimization was initialized with a priori
4
5 52 information.
6
7

8 53 Figure 17. MT responses and error bars for TE and TM modes of three representative stations (12,
9
10 54 13, and 14) of the COPROD2 data set. They show the high quality of the data. The ρ_{app} stands for the apparent
11
12 55 resistivity.
13
14

15 56 Figure 18. L-curve response for COPROD2 data along the horizontal (black diamonds) and vertical
16
17 57 (red circles) directions. The tradeoff between data misfit and model norm indicates the best Lagrange
18
19 58 multiplier λ equal to 0.1.
20
21

22 59 Figure 19. Resistivity model of COPROD2 data from PSO computation, after 6000 iterations.
23
24 60 Lagrange multiplier $\lambda = 0.1$.
25
26

27 61 Figure 20. Reference models of COPROD2 data from Jones (1993b). The 20 stations are sorted and
28
29 62 named as in Figure 19. The color scale for the resistivity (ρ) is consistent with Figure 19: white ($\rho < 1 \text{ ohm-m}$),
30
31 63 pink ($\rho = 1 \text{ ohm-m}$), red ($\rho = 10 \text{ ohm-m}$), yellow ($\rho = 100 \text{ ohm-m}$), and green ($\rho = 1000 \text{ ohm-m}$).
32
33

34 64 Figure 21. Fitting curves between observed apparent resistivity (ρ_{app}) and phase, and predicted
35
36 65 responses at selected periods: 56.9 s, 85.3 s, 341.3 s. Observed data include error bars and are marked with
37
38 66 dots for TE and diamonds for TM. Calculated responses are plotted with solid line for TE and dashed line for
39
40 67 TM. The optimization was randomly initialized.
41
42

43 68 Figure 22. Black curves show computation time in hours (left ordinate axis) as a function of the number
44
45 69 of cores exploited for a reference PSO simulation of 150 iterations with a 10000-particle swarm. The right
46
47 70 ordinate axis and blue curves refer to the total runtime speedup with exploited cores increasing. Dotted lines
48
49 71 refers to “shared” parallel environment (workers of a single node), dashed lines to “orte” (workers of different
50
51 72 nodes).
52
53
54
55
56
57
58
59
60

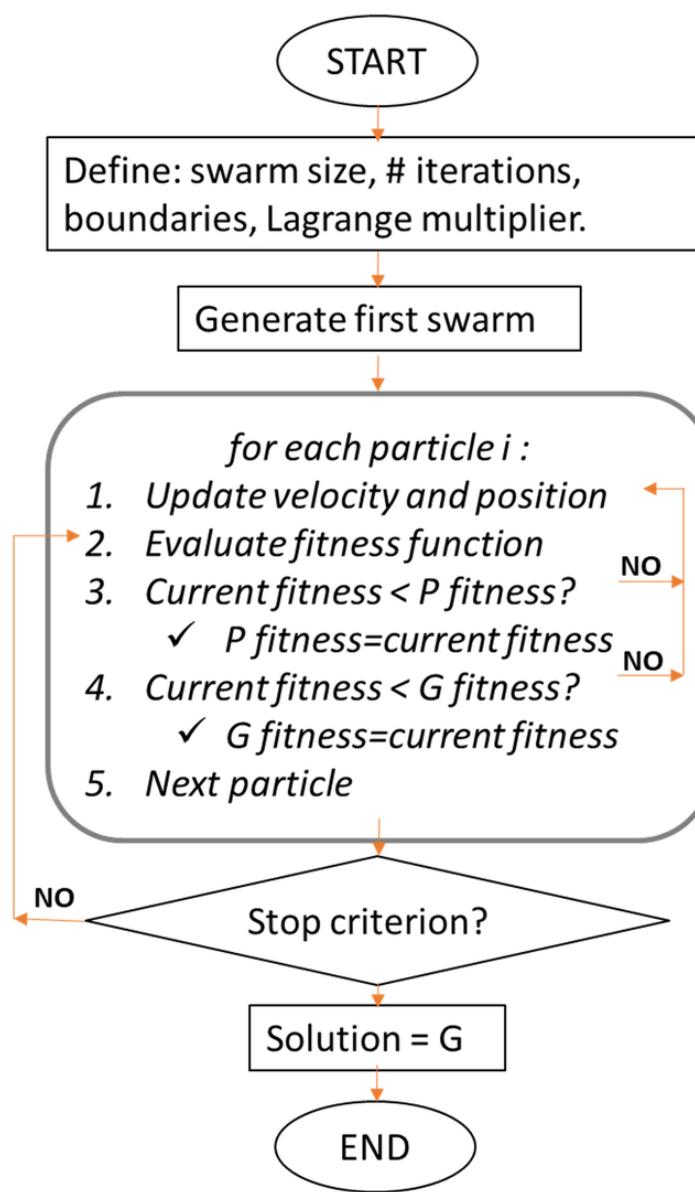


Figure 1. The PSO algorithm flowchart: **P** is the local best solution; **G** is the global best solution.

57x92mm (300 x 300 DPI)

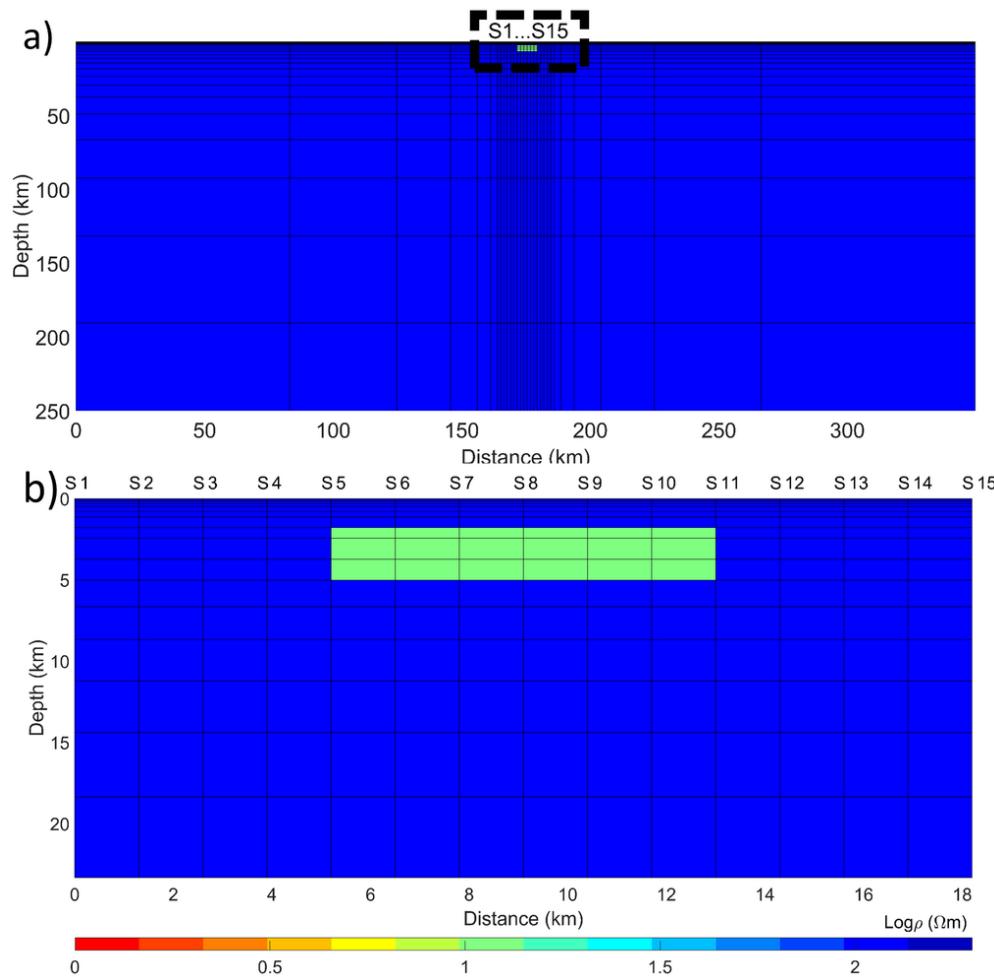


Figure 2ab. Synthetic model 1: a) the 2D mesh is discretized into 33 layers and a total of 957 grid cells. The labels S1, ..., S15 indicate the location of the 15 MT stations. The dashed area is shown in b): A 10 ohm-m conductive body is hosted in a 100 ohm-m medium.

86x85mm (300 x 300 DPI)

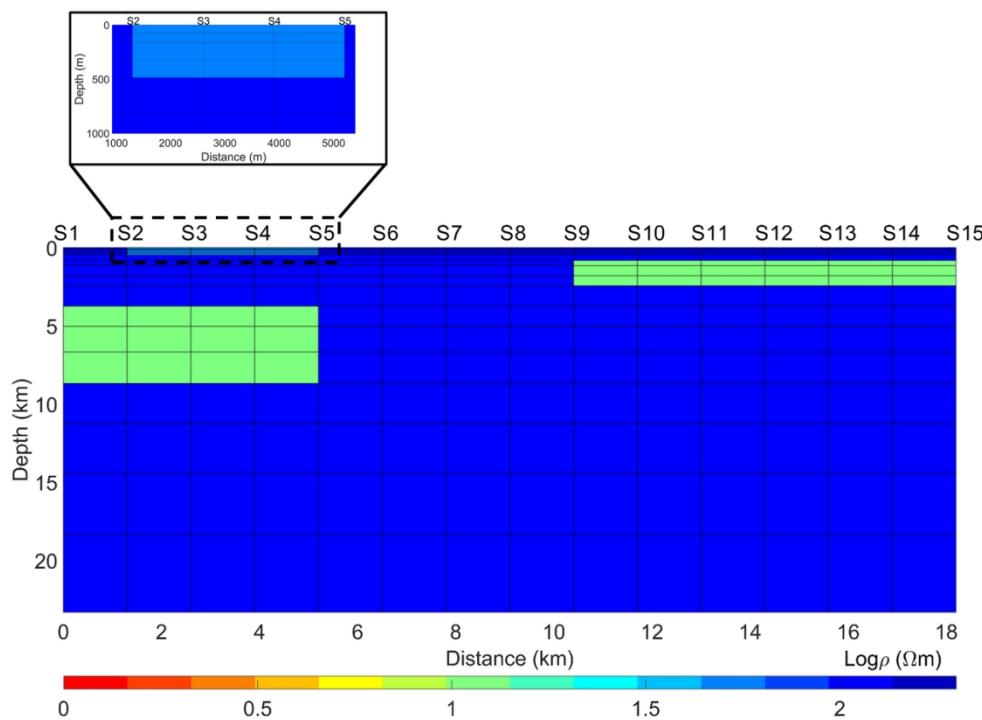


Figure 3. Synthetic model 2: Two 10 ohm-m deep anomalies and one superficial 50 ohm-m body are embedded in a 100 ohm-m host medium. The labels S1, S2, ..., S15 indicate the 15 MT stations. The zoomed-in box on the top shows the 50-ohm-m body below S2-S5.

111x80mm (300 x 300 DPI)

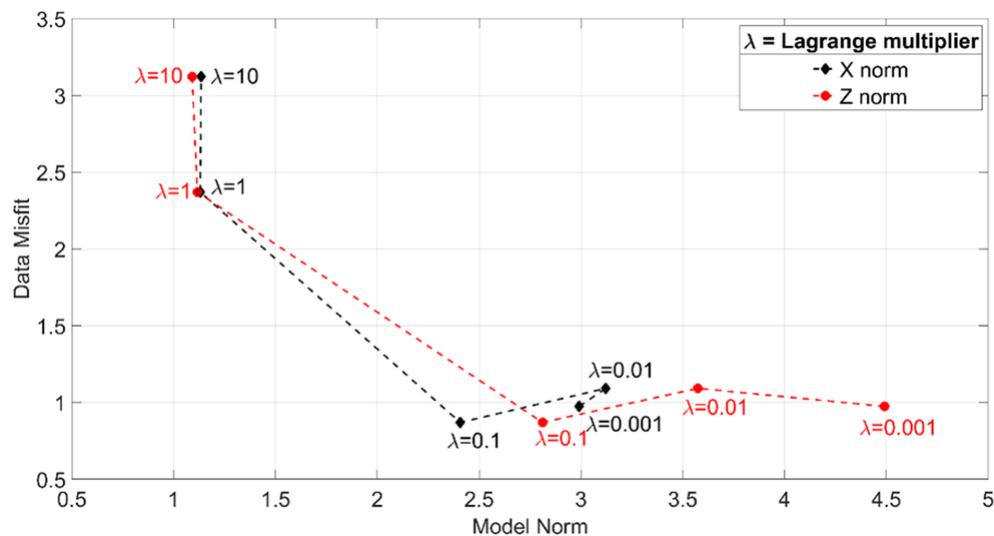


Figure 4. L-curve response for synthetic model 1 along the horizontal (black diamonds) and vertical (red circles) directions. The tradeoff between data misfit and model norm indicates the best Lagrange multiplier λ equal to 0.1.

84x46mm (300 x 300 DPI)

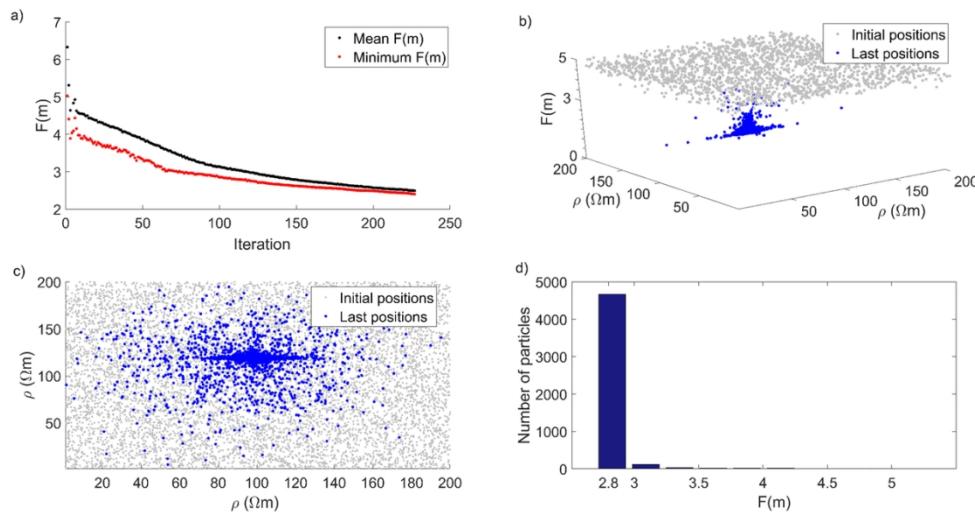


Figure 5. Fitness function $F(\mathbf{m})$ and particle positions at the end of the optimization: a) fitness-function value, iteration after iteration, for the best particle (red dots) and the rest of the swarm (black dots); b) the fitness-function value as a function of the particle positions in the resistivity (ρ) search space, at the first (grey dots) and final (blue dots) iterations; c) plain view of b); d) final distribution of the fitness-function values among all the particles.

110x58mm (300 x 300 DPI)

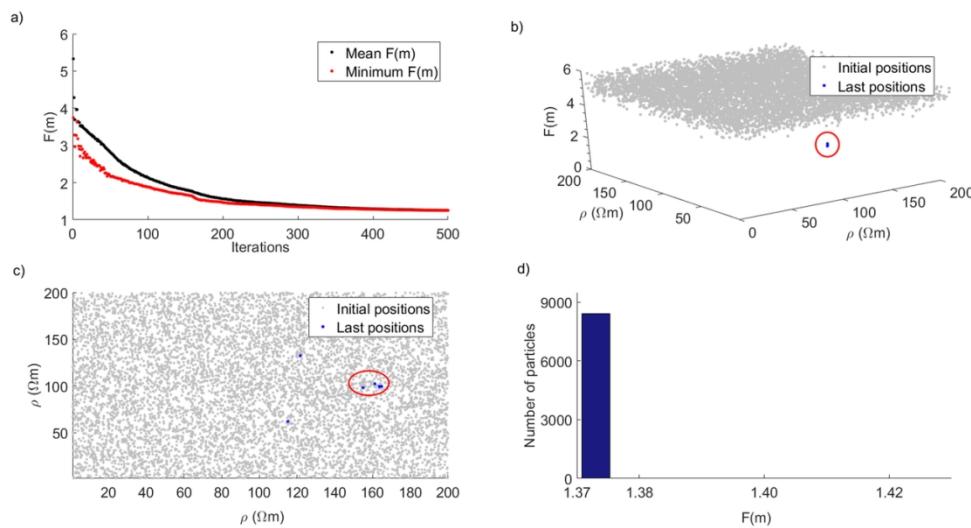


Figure 6. Fitness function $F(\mathbf{m})$ and particle positions at the end of the optimization: a) fitness-function value, iteration after iteration, for the best particle (red dots) and the rest of the swarm (black dots); b) the fitness-function value as a function of the particle positions in the resistivity (ρ) search space, at the first (grey dots) and final (blue dot) iterations; c) plain view of b) with all particles converged to the last position (red circled blue dot); d) final distribution of the fitness-function values among all the particles.

113x60mm (300 x 300 DPI)

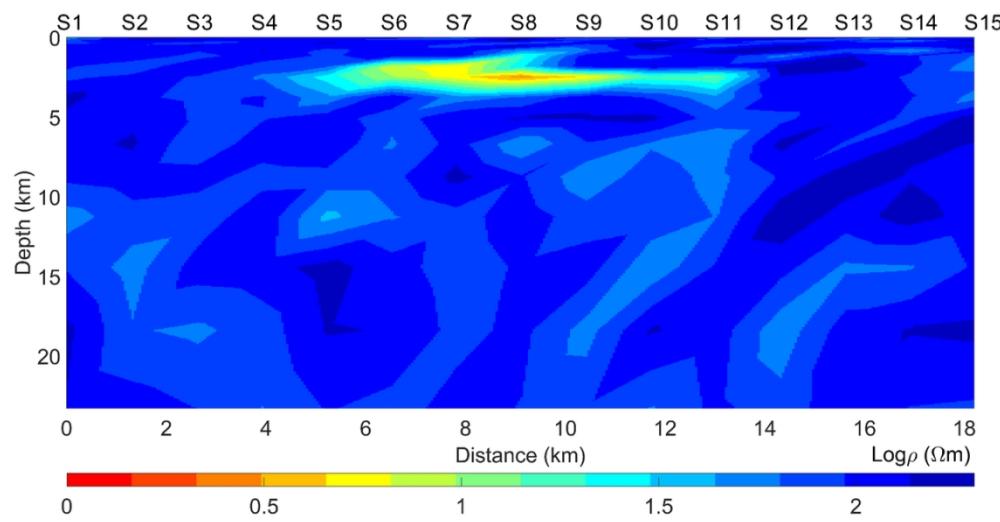


Figure 7. PSO solution for synthetic model 1, after about 150 iterations without a priori initialization for the 8600 particles of the swarm. Lagrange multiplier $\lambda = 0.1$.

109x56mm (300 x 300 DPI)

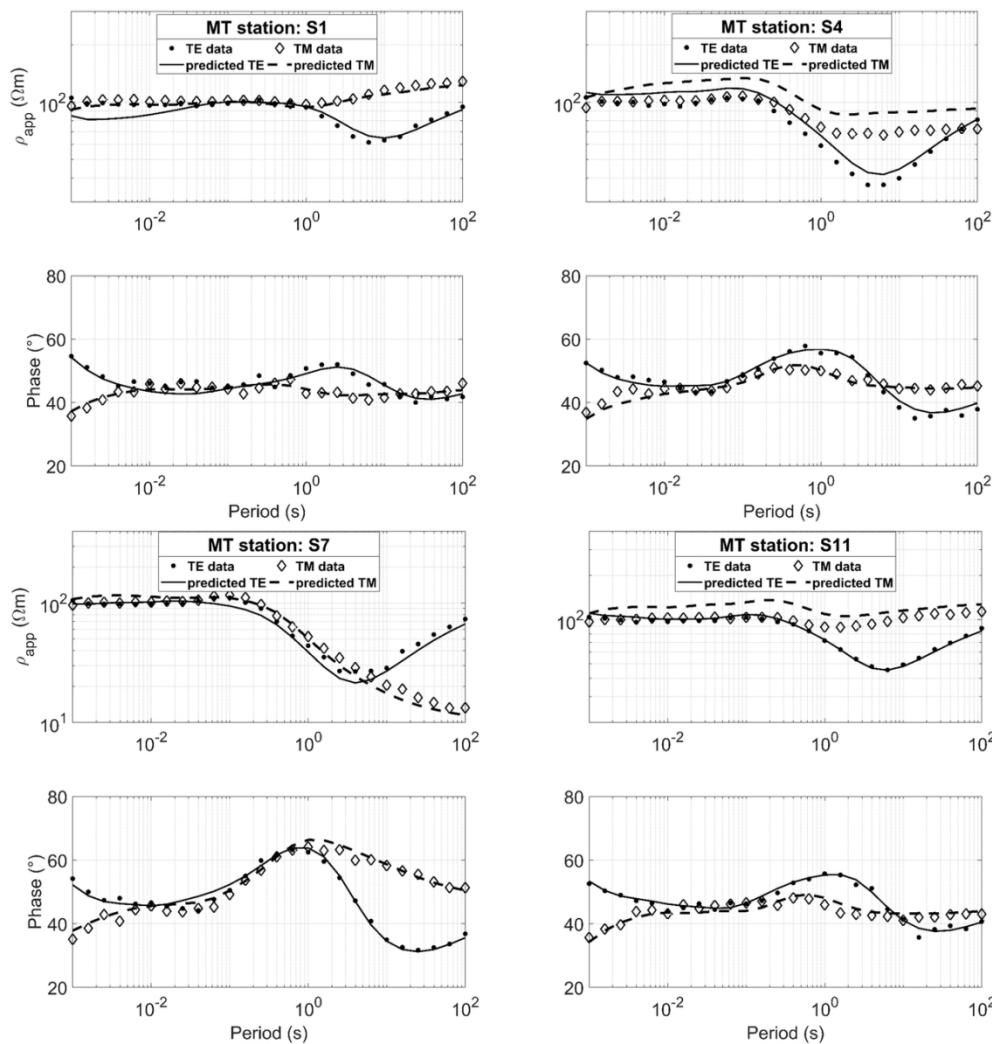


Figure 8. Fitting curves between data of synthetic model 1 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are S1, S4, S7, and S11.

The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are plotted as solid lines for TE, and dashed lines for TM. The optimization was randomly initialized.

109x115mm (300 x 300 DPI)

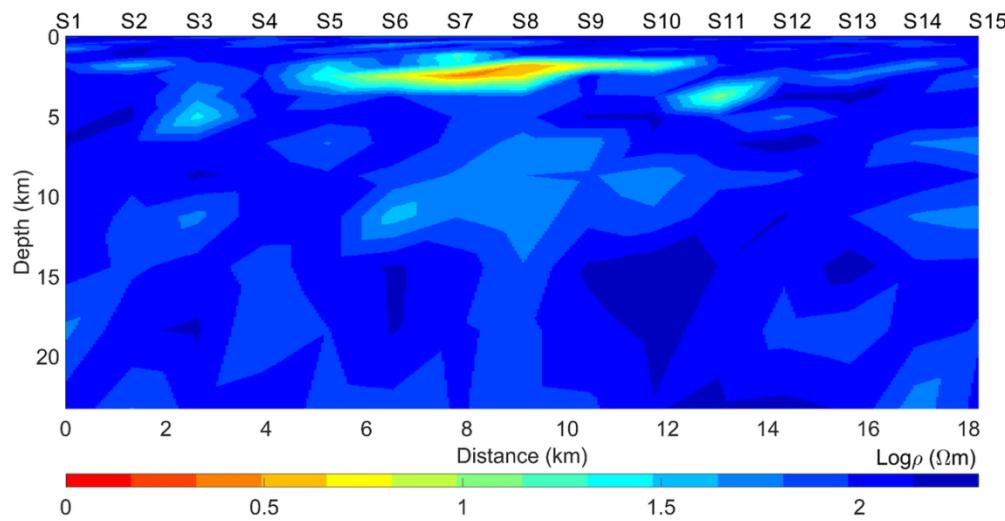


Figure 9. PSO solution for synthetic model 1 using a swarm size of only 5700 particles (6 times the unknowns), after about 160 iterations, without a priori initialization, and Lagrange multiplier $\lambda = 0.1$.

109x58mm (300 x 300 DPI)

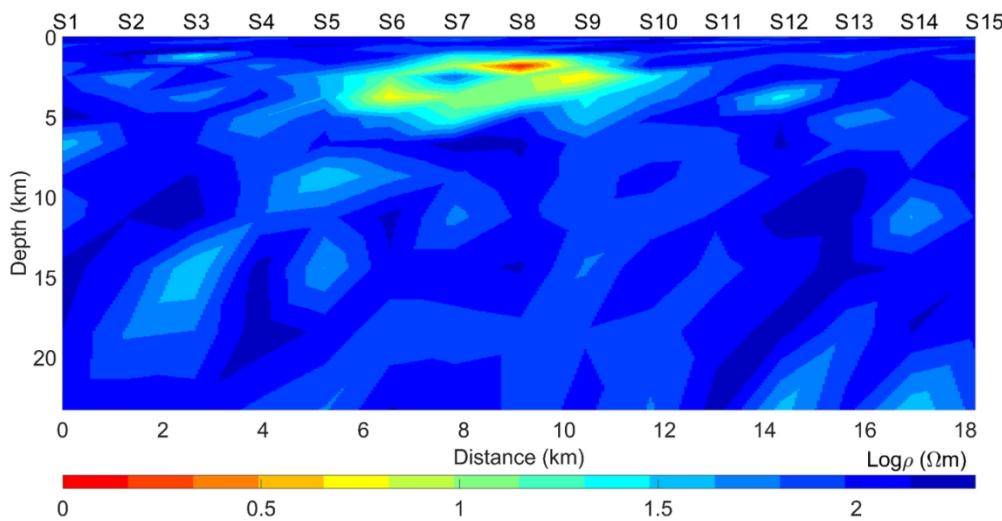


Figure 10. PSO solution for synthetic model 1, after about 250 iterations and with a priori information given to 5% of the particles. Lagrange multiplier $\lambda = 0.1$.

109x58mm (300 x 300 DPI)

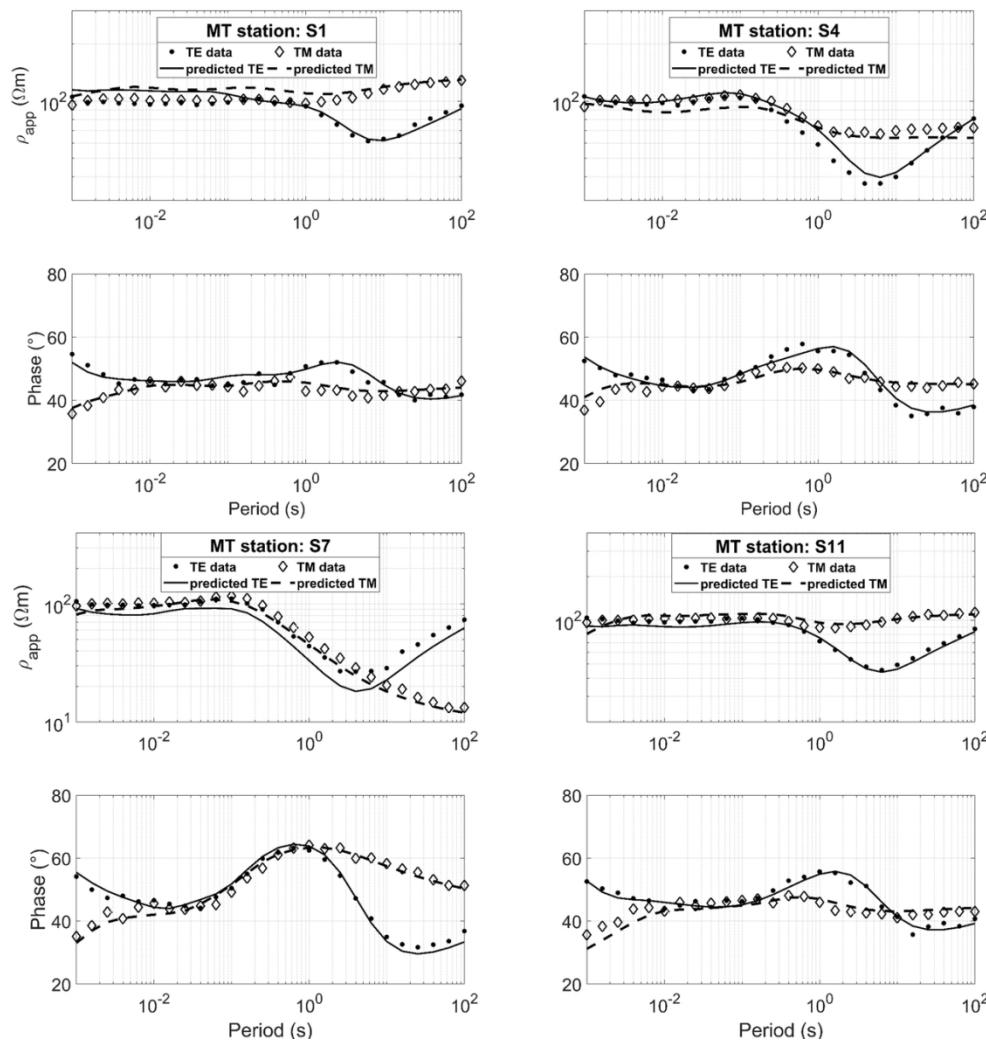


Figure 11. Fitting curves between data of synthetic model 1 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are S1, S4, S7, and S11. The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are plotted as solid lines for TE, and dashed lines for TM. The optimization was initialized with a priori information.

111x115mm (300 x 300 DPI)

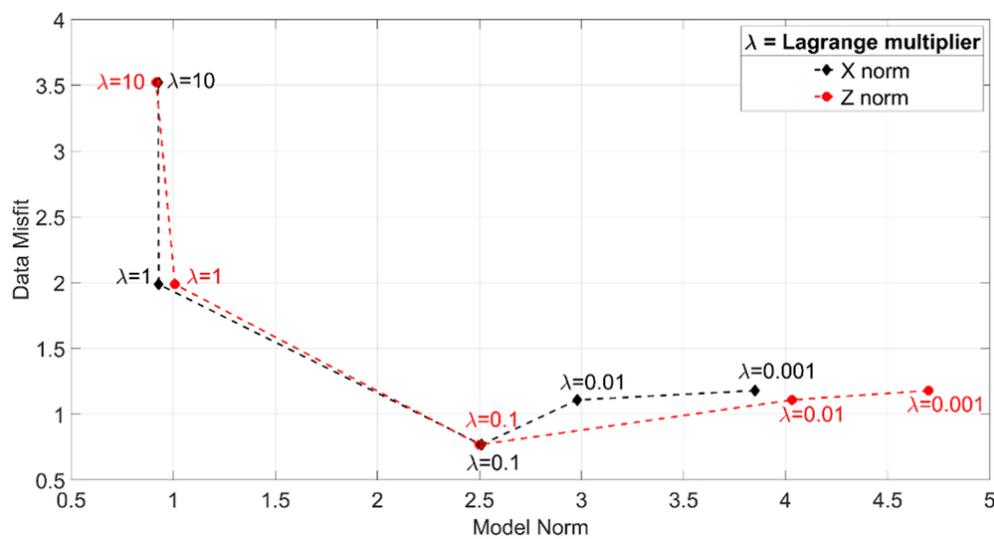


Figure 12. L-curve response for synthetic model 2 along horizontal (black diamonds) and vertical (red circles) directions. The tradeoff between data misfit and model norm indicates the best Lagrange multiplier λ equal to 0.1.

85x45mm (300 x 300 DPI)

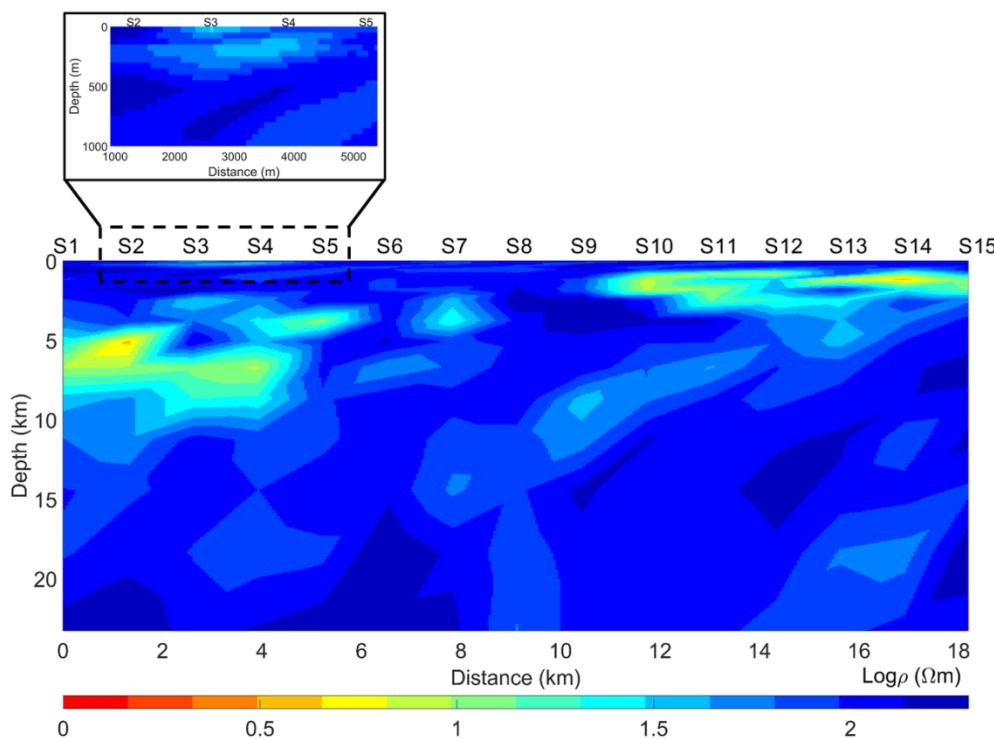


Figure 13. PSO solution for synthetic model 2, after 1674 iterations and without a priori initialization.
Lagrange multiplier $\lambda=0.1$.

109x81mm (300 x 300 DPI)

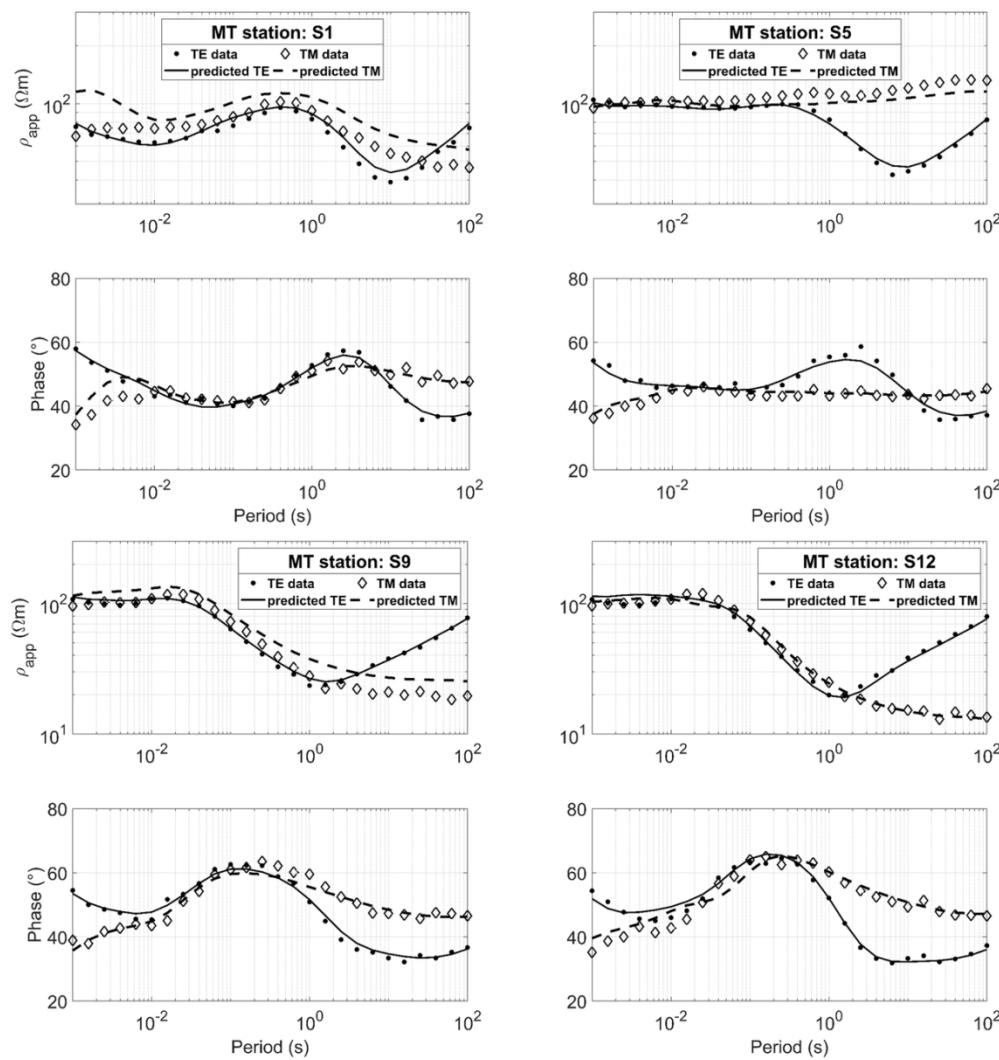


Figure 14. Fitting curves between data of synthetic model 2 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are S1, S5, S9, and S12. The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are plotted as solid lines for TE, and dashed lines for TM. The optimization was randomly initialized.

109x115mm (300 x 300 DPI)

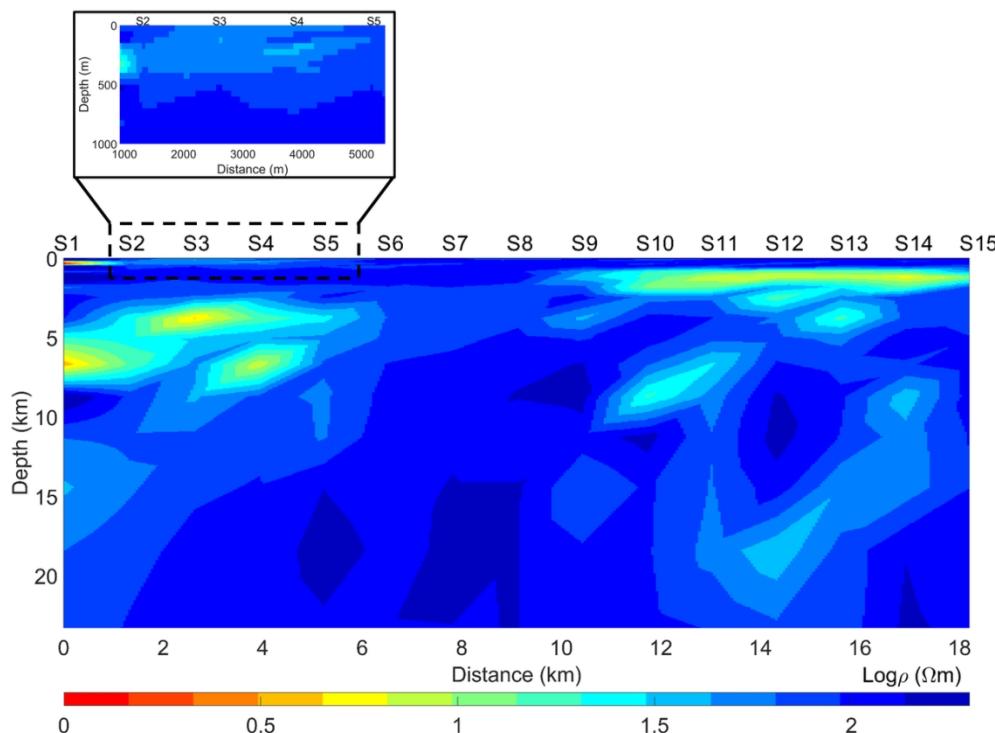


Figure 15. PSO solution for synthetic model 2, after 53 iterations and with a priori information given to 5% of the particles. Lagrange multiplier $\lambda = 0.1$.

110x81mm (300 x 300 DPI)

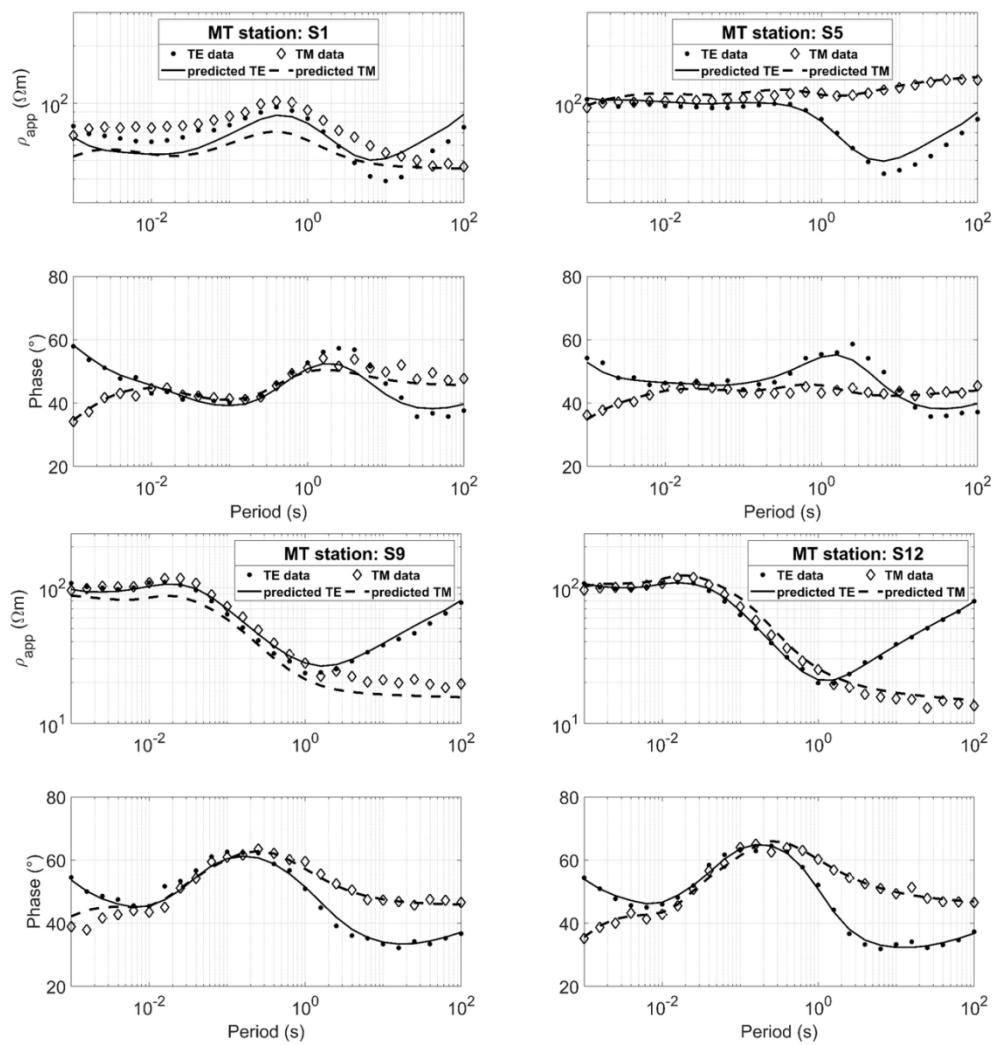


Figure 16. Fitting curves between data of synthetic model 2 and calculated data for apparent resistivity (ρ_{app}) and impedance phase for both TE and TM polarizations. The selected MT stations are S1, S5, S9, and S12. The synthetic data are marked as dots for TE, and diamonds for TM, while the PSO-predicted data are plotted as solid lines for TE, and dashed lines for TM. The optimization was initialized with a priori information.

110x115mm (300 x 300 DPI)

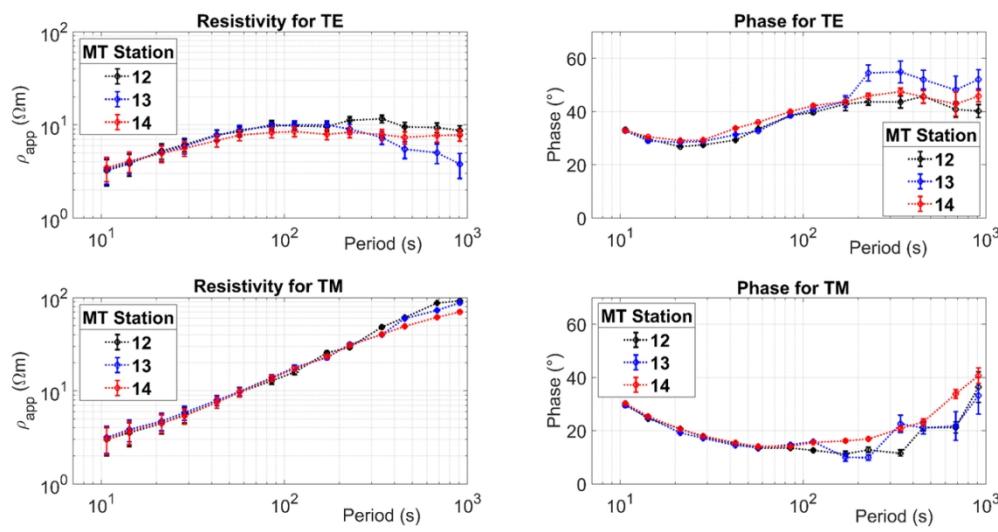


Figure 17. MT responses and error bars for TE and TM modes of three representative stations (12, 13, and 14) of the COPROD2 data set. They show the high quality of the data. The ρ_{app} stands for the apparent resistivity.

108x57mm (300 x 300 DPI)

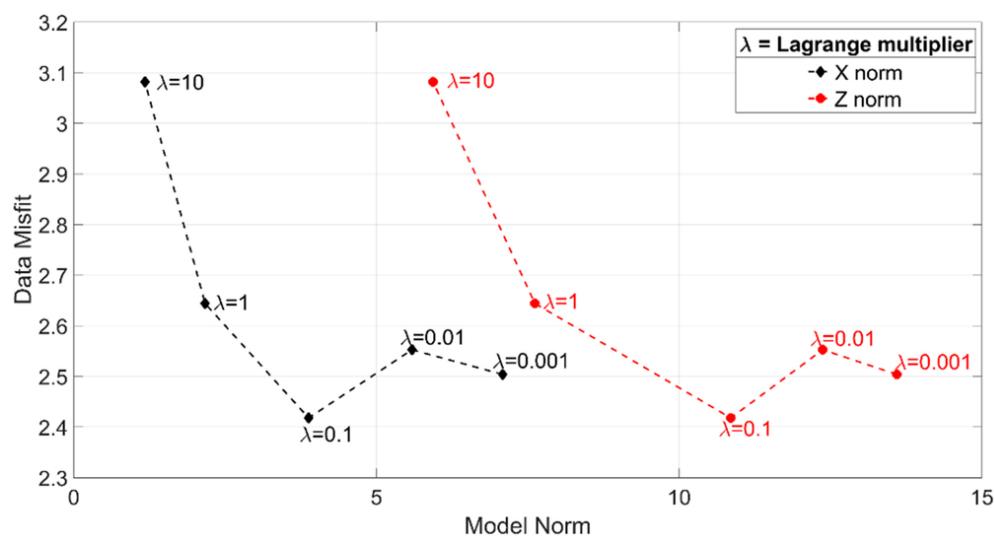


Figure 18. L-curve response for COPROD2 data along the horizontal (black diamonds) and vertical (red circles) directions. The tradeoff between data misfit and model norm indicates the best Lagrange multiplier λ equal to 0.1.

85x45mm (300 x 300 DPI)

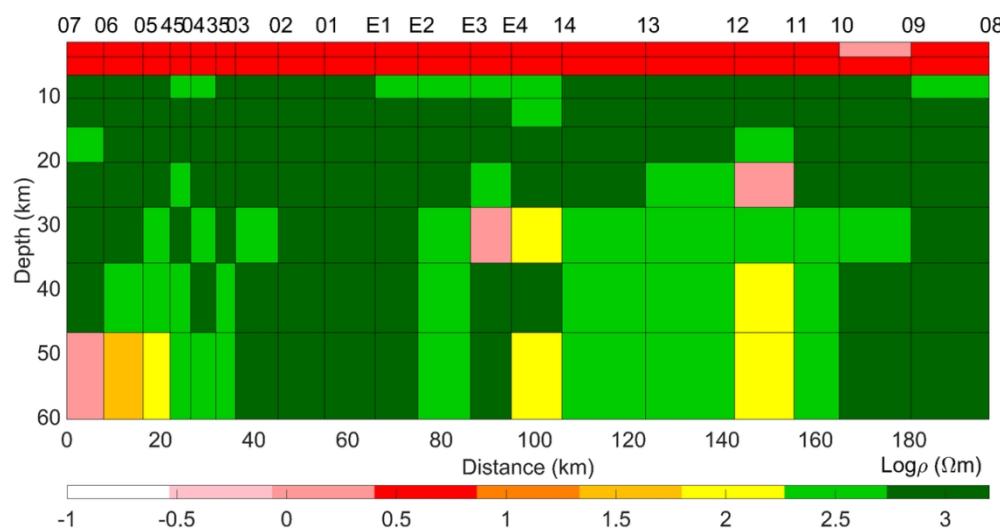


Figure 19. Resistivity model of COPROD2 data from PSO computation, after 6000 iterations. Lagrange multiplier $\lambda = 0.1$.

108x57mm (300 x 300 DPI)

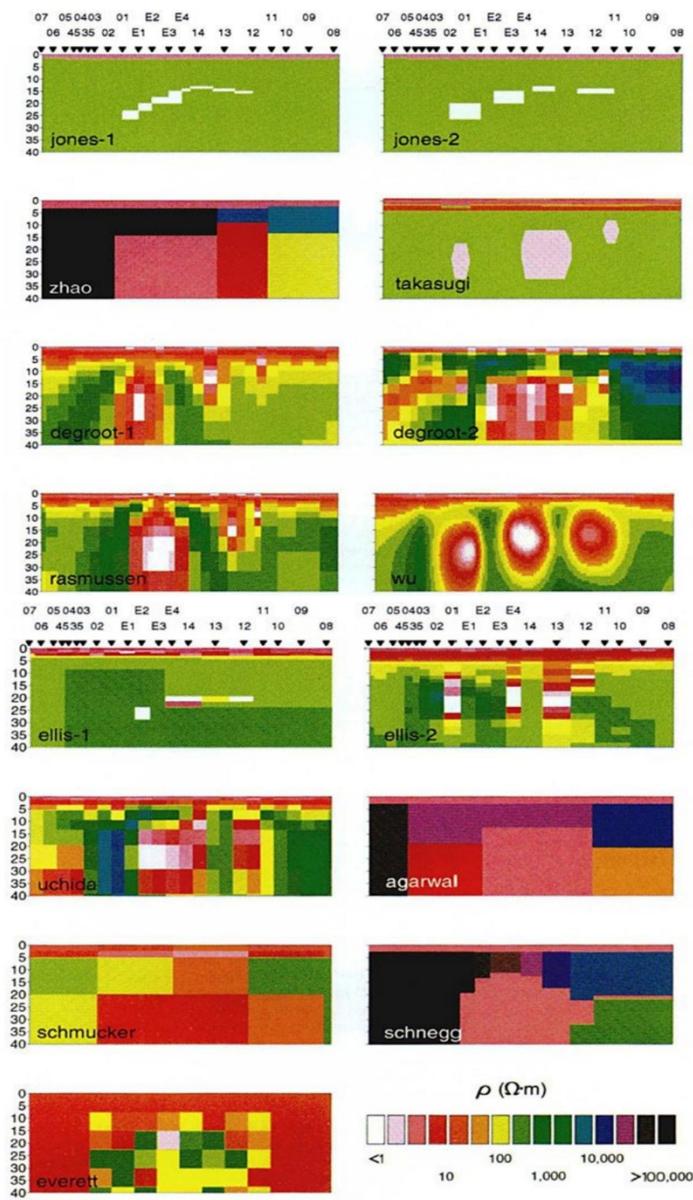


Figure 20. Reference models of COPROD2 data from Jones (1993b). The 20 stations are sorted and named as in Figure 19. The color scale for the resistivity (ρ) is consistent with Figure 19: white ($\rho < 1$ ohm-m), pink ($\rho = 1$ ohm-m), red ($\rho = 10$ ohm-m), yellow ($\rho = 100$ ohm-m), and green ($\rho = 1000$ ohm-m).

110x183mm (300 x 300 DPI)

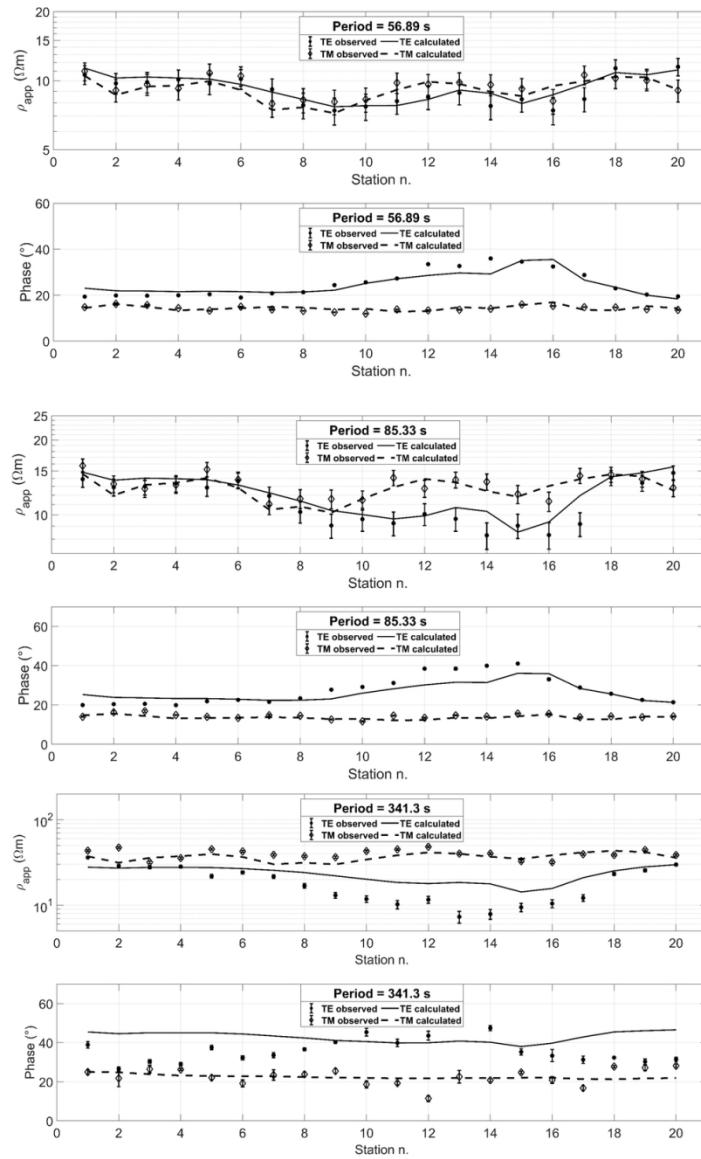


Figure 21. Fitting curves between observed apparent resistivity (ρ_{app}) and phase, and predicted responses at selected periods: 56.9 s, 85.3 s, 341.3 s. Observed data include error bars and are marked with dots for TE and diamonds for TM. Calculated responses are plotted with solid line for TE and dashed line for TM. The optimization was randomly initialized.

87x143mm (300 x 300 DPI)

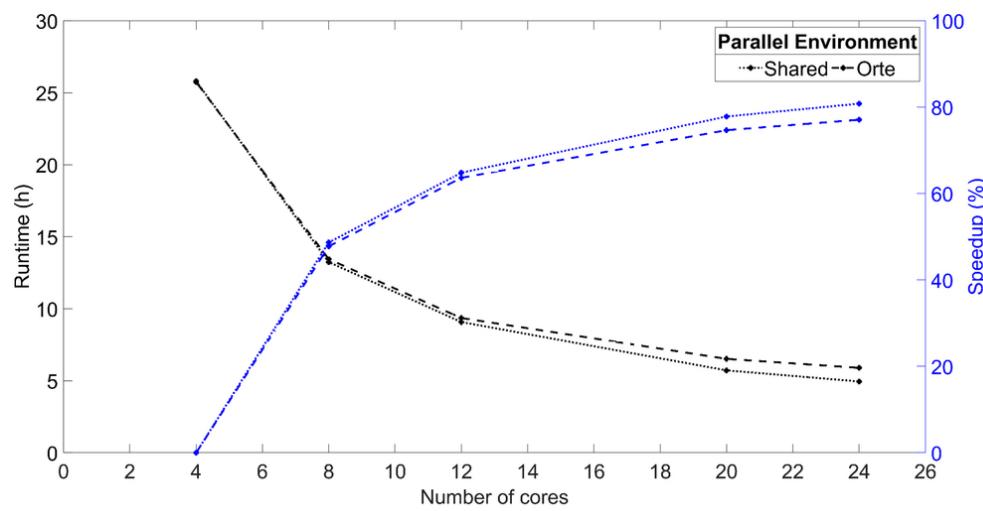


Figure 22. Black curves show computation time in hours (left ordinate axis) as a function of the number of cores exploited for a reference PSO simulation of 150 iterations with a 10000-particle swarm. The right ordinate axis and blue curves refer to the total runtime speedup with exploited cores increasing. Dotted lines refers to "shared" parallel environment (workers of a single node), dashed lines to "orte" (workers of different nodes).

85x43mm (300 x 300 DPI)

LIST OF TABLES

Table 1. Synthetic data from example 1 were adopted to perform the calibration of the cognitive acceleration α_1 and social acceleration α_2 starting from different values at the first iteration ($k=1$). The final values of the RMS error and fitness function $F(\mathbf{m})$ are listed for each test.

| | | $\alpha_2^{k=1} = 0.25$ | $\alpha_2^{k=1} = 0.5$ | $\alpha_2^{k=1} = 0.75$ |
|-------------------------|-----------------------------------|-------------------------|------------------------|-------------------------|
| $\alpha_1^{k=1} = 2.75$ | RMS | 0.91 | 2.17 | 2.18 |
| | $F(\mathbf{m})$ | 1.34 | 2.88 | 3.03 |
| $\alpha_1^{k=1} = 2$ | RMS | 0.88 | 0.86 | 0.91 |
| | $F(\mathbf{m})$ | 1.47 | 1.37 | 1.33 |
| $\alpha_1^{k=1} = 1.5$ | RMS | 1.11 | 1.01 | 0.87 |
| | $F(\mathbf{m})$ | 1.73 | 1.52 | 1.44 |

Table 2. Sensitivity analysis on the population size as PSO input argument. The number of particles was 6, 8, 9, 10, and 12 times the number of unknowns of the problem (957 grid cells). Results are analyzed in terms of: RMS errors, total runtime in hours, and the maximum number of iterations reached.

| Number of particles | Times the unknowns | RMS | Runtime (h) | Iterations |
|----------------------------|---------------------------|------------|--------------------|-------------------|
| 5700 | 6 | 0.88 | 3.47 | 166 |
| 7500 | 8 | 0.90 | 7.17 | 275 |
| 8600 | 9 | 0.86 | 4.60 | 154 |
| 9500 | 10 | 0.88 | 5.82 | 176 |
| 11500 | 12 | 0.87 | 6.52 | 165 |

1
2
3 11 *Table 3. Results of PSO applied to the two synthetic models (with and without a priori initialization)*
4
5 12 *and to the COPROD2 data set (without a priori initialization). Results are presented in terms of:*
6
7 13 *RMS error, runtime, and number of iterations performed before the optimization stop. The runtime is*
8
9 14 *in hours and refers approximately to one single trial.*

| Data set | Initialization | RMS | Runtime (h) | Iterations |
|--------------------------|-----------------------|------------|--------------------|-------------------|
| <i>Synthetic model 1</i> | <i>No a priori</i> | 0.86 | 4.6 | 154 |
| | <i>A priori</i> | 0.91 | 3.17 | 250 |
| <i>Synthetic model 2</i> | <i>No a priori</i> | 0.9 | 28.8 | 1674 |
| | <i>A priori</i> | 0.99 | 0.55 | 53 |
| <i>COPROD2</i> | <i>No a priori</i> | 2.42 | 8 | 6000 |

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.