The design of an open-source carbonate reservoir model

Jorge Costa Gomes*, Sebastian Geiger and Daniel Arnold

Institute of GeoEnergy Engineering, Heriot-Watt University, UK

© 2022 The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License (http://creativecommons.org/licenses/by/4.0/). Published by The Geological Society of London for GSL and EAGE. Publishing disclaimer: www.geolsoc.org.uk/pub_ethics

Abstract: This work presents a new open-source carbonate reservoir case study, the COSTA model, that uniquely considers significant uncertainties inherent to carbonate reservoirs, providing a far more challenging and realistic benchmarking test for a range of geo-energy applications. The COSTA field is large, with many wells and large associated volumes. The dataset embeds many interacting geological and petrophysical uncertainties in an ensemble of model concepts with realistic geological and model complexity levels and varying production profiles. The resulting number of different models and long run times creates a more demanding computational challenge than current benchmarking models.

The COSTA model takes inspiration from the shelf-to-basin geological setting of the Upper Kharib Member (Early Cretaceous), one of the most prolific aggradational parasequence carbonate formations sets in the world. The dataset is based on 43 wells and the corresponding fully anonymised data from the northeastern part of the Rub Al Khali basin, a sub-basin of the wider Arabian Basin. Our model encapsulates the large-scale geological setting and reservoir heterogeneities found across the shelf-to-basin profile, into one single model, for geological modelling and reservoir simulation studies.

The result of this research is a semi-synthetic but geologically realistic suite of carbonate reservoir models that capture a wide range of geological, petrophysical, and geomodelling uncertainties and that can be history-matched against an undisclosed, synthetic ‘truth case’. The models and dataset are made available as open-source to analyse several issues related to testing new numerical algorithms for geological modelling, uncertainty quantification, reservoir simulation, history matching, optimization and machine learning.

Supplementary material: supplementary data associated with this article can be found in the data repository for the COSTA model available at https://doi.org/10.17661/6e36e28d-50d9-4e31-9790-18db4bce6e5d and supplementary material is available at https://doi.org/10.6084/m9.figshare.c.5823571

Received 16 August 2021; revised 7 January 2022; accepted 2 February 2022

Researchers need benchmarking examples to test and validate their reservoir modelling techniques based on simulation speed or accuracy, complex development optimization and history matching. Current open-source offerings are predominantly from clastic reservoirs and often geologically simplified so that they are fast to run (e.g. the PUNQ S3, SPE10, Brugge Field, Nome Field models or the Volve Field) (Cutts 1991; Christie and Blunt 2001; Castro et al. 2005; Statoil 2005; Equinor 2018). Many of these datasets provide just one geological scenario, ignoring geological uncertainties or simplifying them considerably.

For clastic reservoirs, the Watt field was developed to provide an ensemble of geological models that covers a wide range of typical uncertainties (e.g. formation tops, fault location, petrophysical properties and facies distributions) and modelling choices (e.g. permeability cut-offs, pixel v. object-based modelling) (Arnold et al. 2013). The Watt Field created a uniquely challenging history matching benchmark given the widely varying performance of the provided geological models, which did not include the ‘truth case’ model.

For carbonates, the most recent open-source models are UNISIM-II-D, UNISIM-II-M, UNISIM-II-H, and UNISIM-III (Correia et al. 2015, 2020). These models have a range of applications, such as field development, optimization and reservoir management. UNISIM-II-H captures uncertainties in an ensemble of property realizations, such as porosity and permeability fields. While these realizations capture some of the petrophysical uncertainties inherent to carbonate reservoirs, they only encompass a very narrow range of uncertainties. More fundamental uncertainties arising from interpreting the available data and developing geological concepts in different ways have not been considered. Therefore, we need a new benchmark study for complex carbonate reservoirs that captures the inherent geological uncertainties and accounts for the large model sizes needed when modelling carbonate reservoirs.

The novelty of the COSTA model is that it provides a unique dataset that captures an extensive range of geological, petrophysical, and geomodelling uncertainties, which are typical for an aggradational carbonate build-up reservoir. We captured uncertainties by developing multiple geological interpretations, creating different reservoir architectures that behave dynamically differently (Fig. 1). This approach better captures geological and interpretational uncertainties than relying on stochastic approach but requires more effort from the geomodeller (Ringrose and Bentley 2015; Bentley 2016). We include geological uncertainties hierarchically as an ensemble of models, from the interpretation of the top structures to the stratigraphic framework (i.e. the reservoir and baffle zone tops and thicknesses), the facies interpretation, the diagenetic overprints, the reservoir flow unit architecture (i.e. cumulative flow and storage capacity), the wettability variations, and the free water level (FWL).

The second novelty of this work has to do with the origin of the dataset and the model philosophy. We captured a downsized shelf-to-basin geological setting of the Upper Kharib Member into one single model with three main anticlines. We wanted to encapsulate different carbonate platform characteristics and reservoir flow unit behaviours for various geomodelling and reservoir simulation studies. We used fully anonymised data from 43 wells from several fields belonging to different depositional settings and heights above the FWL within an area of 36 000 km². The data comprises a full...
suite of open-hole logs and core data that has been anonymised, rescaled, repositioned, and structurally deformed; FWLs were normalized, and the entire model was placed in a unique coordinate system. The sub-layering and zonation in the models reflect the actual large-scale shelf-to-basin reservoir characteristics and architectures of the Upper Kharaib Member but in an entirely new synthetic geological structure (Alsharhan 1993; Oswald et al. 1995; Grotsch et al. 1998; Davies et al. 2002; Strohmenger et al. 2004, 2006; Melville et al. 2005; Al Mansoori et al. 2008; Buchem et al. 2010a, b; Maalouf et al. 2016; Jingjing et al. 2018; Ferran Pacheco et al. 2019).

This paper is split into three main components: summarizing our undisclosed ‘truth case’ model, the design of 144 open-source model variants, and the generation of 30 years of synthetic production data. We also showcase the effect of grid refinement on our model along with the envelope of uncertainty in dynamic reservoir performance (e.g. production profiles and water-cut evolution) of all model releases. Supplementary data associated with this article can be found in the data repository for the COSTA model at https://doi.org/10.17861/6e36e28d-50d9-4e31-9790-18db4bec6e5d.

**Dataset overview**

The COSTA dataset contains measured field data, static and dynamic models, and production data.

The well dataset comprises a mixture of real and synthetic data. We provide synthetic contour maps, well coordinates, a special core analysis (SCAL) database, fluid properties, capillary pressure (drainage and imbibition), and relative permeability curves for reservoir simulation. We based the open-hole log and core data on actual field data.

We provide 72 model scenarios encapsulating 14 different uncertainties. We kept all the spatial geostatistical parameters of the variograms (e.g. nugget, sill, range, azimuth, variogram models) consistent throughout the modelling of all scenarios. However, we re-modelled all 72 original model scenarios with a reduced spatial variogram range for porosity and permeability. In total, 144 models are available as part of the open-source package.

Included are 30 years of synthetic production data generated from our undisclosed ‘truth case’ model to obtain field-wide and well-by-well production data (oil, gas, water rates, bottom-hole pressures, etc.) history matching. The production originates from a synthetic field development plan (FDP) devised over 15 years. The FDP was designed to mimic an onshore drilling campaign typical for giant onshore carbonate reservoirs in the Middle East. We added random Gaussian noise of 5% to the production data. To the bottom hole pressures (BHPs), we added 1% to simulate errors in production measurements.

We created the first ensemble to select one model to act as the ‘truth case’, but only one truth case (or true model) exists. We then created two separate ensembles of models for release that honour the data but are different from the ‘truth case’. Figure 1 explains that our undisclosed ‘truth case’ has a unique stock tank original oil-in-place, stratigraphic framework, reservoir flow unit architecture, facies distribution, 27 reservoir rock types, wettability, multiple oil–water contacts and capillary transition zone thicknesses, free water level and skin effect (±5) on all producing wells.

**Geological concept**

We created three ensembles: an undisclosed ‘truth case ensemble’ from which we selected one ensemble member to act as our undisclosed ‘truth case’, ensemble 1, and ensemble 2 (Fig. 1). Each
ensemble has its unique geological concept of the stratigraphic framework and unique reservoir flow unit architecture. We kept the number of layers and reservoir and baffle zonation's constant. The Net-to-Gross varies between 80–90%.

**Choice of data**

Each ensemble has a choice of wireline and cored porosity well data and models. Porosity from logs is generally measured up to over 1000 times the core plug volume, and measuring points vary between, i.e. 0.3–1 ft. intervals. Therefore, significant uncertainties can be inherent in the porosity modelling (Rider 1996). In this study, the pore volume difference between cored (higher) and logged (lower) porosity values are 3%, and hence each contributes to a unique stock tank original oil-in-place (STOIP).

Our synthetic SCAL database provides the modeller with many water saturation profiles above the FWL for model initialisation. It contains 110 capillary pressure and pore throat size distribution (PTSD) curves from anonymised Mercury Injection Capillary Pressure (MICP) carbonate plug samples linked with porosity, permeability, and Winland R35 values.

**Modelling choices**

Each ensemble has nine permeability maps, six bivariate, and three non-bivariate distributions. The spatial bivariate permeability distributions were modelled linked with either wireline or cored porosity values (and facies). We also include individual permeability distributions, which use either arithmetic, geometric, or harmonic vertical permeability averaging. Furthermore, three non-bivariate permeability distribution maps are included because permeability prediction at the well level was not an issue.

**Model complexity**

Each ensemble has six uniquely generated carbonate build-up scenarios (facies maps) for each reservoir layer. We initially modelled 27 facies distributions for the ‘truth case’ and later discretised the facies into 3, 5, and 7 reservoir rock type (RRT) model variants. We clustered the facies into smaller rock types to reduce the model complexity and influence the sweep and drainage during reservoir simulation.

**Model anisotropy**

Seventy-two model variants were initially created and later remodelled using a reduced anisotropic variogram range for porosity and permeability. We reduced the anisotropic spatial range to see the impact on field performance.

**Summary of uncertainties**

We incorporated a total of 14 uncertainties in our ensemble, all of which will be elaborated in this paper. The uncertainties cover complexities that affect the microscopic displacement efficiency (i.e. capillary pressure, relative permeability, wettability), areal sweep efficiency (i.e. spatial heterogeneity, permeability anisotropy, facies distribution), and vertical sweep efficiency (i.e. stratigraphic layering, flow units, vertical heterogeneity, hydrocarbon columns) during field development. We did not include fractures nor faults in this study, and we kept the PVT consistent in all ensemble members.

(1) Stratigraphic framework
(2) Net-to-Gross
(3) Facies distribution
(4) Diagenetic overprint
(5) Number of RRTs
(6) Porosity source (wireline v. cored)

(7) Vertical permeability profile (arithmetic, geometric & harmonic averaging)
(8) Spatial permeability distribution (bivariate v. non-bivariate)
(9) Spatial variability (variogram) petrophysical property distribution
(10) STOIP
(11) FWL
(12) Oil-water contacts (OWCs)
(13) Capillary transition zone thickness (TZ)
(14) Formation wettability

We also included production-related uncertainties:

(1) Undisclosed skin effect (±5) on all producer wells belonging to our ‘truth case’.
(2) Zero skin effect on all producer wells belonging to open-source ensemble members.
(3) Random Gaussian noise added to oil, water, and gas rates (5%).
(4) Random Gaussian noise added to BHP (1%).

The 144 models that are part of the open-source dataset are not history-matched. These models will provide the reservoir simulation engineers with the opportunity to history-match them against the 30 years of synthetic production from our undisclosed ‘truth case’.

**Geological analogue – Upper Kharaib Member**

Our open-source reservoir models are a realistic and geologically consistent analogue for the aggradational parasequence carbonate formation characteristics from the Upper Kharaib Member (Early Cretaceous) found in and around the Rub Al Khali basin in the Middle East. The Upper Kharaib Member is being produced in many onshore and offshore fields of the basin and has diverse geological flow unit architectures and petrophysical property distributions (Buchem et al. 2010a, b). Moreover, this geological formation exhibits a significant heterogeneity contrast between upper and lower zones. The upper zone is mainly dominated by grain-supported limestone fabrics generally subjected to early diagenetic dissolution, allowing for high permeability (100 mD to 1D) (Carvalho et al. 2011). Mud-supported limestones generally dominate the lower section with low permeability throughout (1 to 10 mD) (Carvalho et al. 2011; Jingjing et al. 2018; Ehrenberg et al. 2020a, b). The formation is roughly 160 ft. thick, and its thickness remains relatively constant across the region. To fully appreciate the geological scale of this aggradational parasequence formation, it is found in and around the Rub Al Khali basin across multiple countries in the Middle East (Fig. 2). Within the U.A.E., it is estimated to stretch across an area of 36,000 km². Due to the subsidence of the Arabian plate with the Eurasian Plate, the formation is found in a series of antiformal structural traps which are progressively getting deeper towards the NE (from current day onshore to offshore) (Ali et al. 2017; Al Blooshi 2018).

The philosophy of this work was to attempt to convert the large-scale geological features and reservoir heterogeneities present in the main fields producing from the Upper Kharaib Member from shelf-to-basin to a reasonably sized and manageable 3D geo-model (Fig. 3). We wanted to honour the large-scale depositional environments from shelf-to-basin and preserve the formation’s typical large-scale reservoir architecture and sequence stratigraphy. It was essential to capture reservoir heterogeneities (porosity, permeability variations) due to diagenetic overprints, including baffles/tight zones and high permeability streaks seen in the literature. Ultimately, we needed to ensure the typical fluid distributions at each well were honoured.
Overview

In this section, we will describe key aspects of the model design of the COSTA model. Some parts of the model construction are discussed in more detail in Costa Gomes et al. (2019a, b, 2020). The COSTA model is roughly 160 ft thick and consists of 15 geological zones, which have been further subdivided into 62 geological layers. The thickness of the grid cells varies between 0.8 and 8 ft, and all the grid cells are 250 × 250 m long. The layering thickness depends on the detail needed to capture the vertical heterogeneity.

All 144 ensemble members originate from a sector of the full COSTA model shown in Figure 4. This area is assumed to reflect the shelf-to-platform geological characteristics of the Upper Kharaib Member and contains 17 out of the total 43 wells used in its construction. The wireline log and core data from these 17 wells are provided in the open-access dataset for those interested in reconstructing this sector. Structurally, the reservoir is a four-way anticlinal closure roughly 60 km long and 26 km wide with flanks dipping 0.6° in northern and southern regions and 1° in the western and eastern areas. The longer axis of the reservoir is trending NE–SW. All model releases have 1.3 million active grid blocks, typical for some giant carbonate reservoirs in the Middle East.

Geological concept

Instead of a seismic cube, a unique conceptualized model with a synthetic contour map was designed and used. The wireline and core data from several fields located across the shelf-to-basin profile originated from multiple anticlinal structural traps which get deeper towards the basin, i.e. in a northeasterly direction. To remove the local field structures and downsize the estimated area of 36 000 km² down to roughly 8.300 km², the spatial distances of the wells were minimized proportionally. We then set a new common FWL across the entire model. The well depths were adjusted to the new FWL in such a way that they honour the hydrocarbon columns seen in the logs. This process allowed us to restructure the large-scale geological setting and the well locations to our synthetic geological structure (Figs 5 and 6). This method also allowed us, as a whole, to honour the actual stratigraphic palaeo slope (still c. 1°). In summary, the large-scale depositional environments, the typical large-scale reservoir architecture, and the sequence stratigraphy found across the shelf-to-basin profile of the Upper Kharaib Member has been...
represented for the first time into one single model with three main anticlines for reservoir simulation and engineering studies.

**Framework modelling**

While creating the COSTA model, the first uncertainty focused on the framework modelling, namely modelling the surfaces, i.e. surfaces based on well tops v. surfaces based on both well tops and digitised contour maps. Creating surfaces using the combination of well tops and contour maps v. using only one or the other provides a unique structural interpretation map. Surfaces are also influenced by the chosen modelling algorithm (i.e. convergent, least squares, etc.). In a real-life field case study, the structure will inevitably alter throughout time as more wells are drilled. Hence the structural uncertainty and gross rock volume are vital uncertainties that should be updated throughout the life of a field. This study provides two unique contour map interpretations for those willing to rebuild their geological models. However, we chose one representative structural map for the life of the field. Examples are shown in the supplementary material.

**Free water level**

The fluid-petrophysical uncertainty of where to place the FWL was considered in the model design. All 144 ensemble members have been assigned a unique variation in FWL of ±30 ft. The saturation distributions can represent (at large) the saturation derived from the open-hole logs even with a variation in the FWL of ±30 ft. However, changing the FWL ±100 ft (extreme case) has not only a dramatic impact on STOOIP estimates but is also significant for those working on carbon capture and storage projects (Fig. 7). For example, the amount of CO₂ injected and its storage design will be directly influenced by the spill points of the reservoir closures found across the model (Ajayi et al. 2019).

**Synthetic water saturation profiles**

We used real wireline log (i.e. gamma-ray, density, neutron porosity, sonic, effective porosity, resistivity, water saturation) and core data (i.e. effective porosity and absolute permeability) from 43 wells in this study. Although most of the wireline logs comprise a...
full data suite, several wells did not have resistivity. Therefore, we back-calculated resistivity trends from wells with saturation profiles using a Generalised Reduced Gradient (GRG) Non-linear Solver. The key petrophysical uncertainties in Archie’s saturation equation are the exponents $n$ and $m$ values which are not measured downhole. The parameter $n$ is the saturation exponent and a function of wettability, and the parameter $m$ is the cementation exponent, which is a function of tortuosity. Many petrophysicists opt for a value of 2 for each; however, these values can vary significantly in the formation, often between 1.7 to 5. A higher $n$ value would signify a more oil-wet rock and therefore reduce the oil in place for that specific interval. We interpreted our variable $n$ and $m$ parameters with height above the FWL. Saturation profiles are shown in the supplementary material.

Reservoir flow unit architecture

The subsequent fundamental geological uncertainty is the reservoir flow unit architecture based on the stratigraphy and the diagenetic overprints of the formation, which control the porosity and permeability distribution. This study assembled three different flow unit architectures: our undisclosed ‘truth case ensemble’, ensemble 1, and ensemble 2. All ensembles have unique interpretations of the formation tops, baffle tops and their respective thicknesses, along with the cumulative flow and storage capacity for every flow unit. Building three different geological concepts of the stratigraphic framework and flow unit architecture affects the overall vertical profile of petrophysical property distributions. All three flow unit architecture concepts have eight reservoir zones and seven baffle zones with an average Net-to-Gross of 85%.

Figure 8 shows one of three flow unit architecture concepts belonging to the 43 wells used to construct the full COSTA model. The stratigraphic modified Lorenz plots (SMLP) are used to quantitatively assess the dynamic characterization of flow and storage capacity within each reservoir unit. This study mimics natural aggradational carbonate build-up stratal geometries and internal reservoir architectures by preserving baffles and flow units along with realistic porosity and permeability trends observed in the literature (Wu and Ehrenberg 2016; Ehrenberg and Wu 2019; Ehrenberg et al. 2020a, b).
The average Lorenz coefficient is 0.59, and the average porosity and permeability coefficient of variance (Cv) are 0.36 and 1.83, respectively. However, there is a heterogeneous contrast between the upper and lower parts of the model. The upper part of the model has an average Cv of 0.47 for porosity and an average Cv of 1.01 for permeability. The lower part of the model has an average Cv of 0.28 for porosity and 0.89 for permeability. Figure 8 demonstrates via the SMLP and normalized cumulative flow capacity plots that most of the flow originates from the upper part of the model.

**Facies and reservoir rock types**

The following three elements of the model construction are all intertwined and susceptible to uncertainties: facies maps, diagenetic overprints, and reservoir rock typing. Facies maps are designed to conceptualize the subsurface carbonate depositional environment. Typically, facies are identified through detailed sedimentological studies and are mapped in 2D. However, envisaging a realistic carbonate depositional environment is challenging because carbonate petrophysical properties vary from those of modern sediments due to multiple diagenetic overprints that can occur throughout burial history. Diagenesis can reduce porosity, alter permeability by redistribution of pore space (dolomitization), and improve the rock’s petrophysical quality (Lucia 1999). The impact of diagenetic overprints on original fabrics needs to be captured in reservoir modelling workflows by digitising them into the modelling software. These overprints are then used for trending petrophysical properties such as porosity and permeability in 3D space.

We only core a fraction of the total wells in a reservoir in the real world. Therefore, many missing data at the well level, i.e. core permeability, cannot be predicted without a proper reservoir rock typing methodology which is linked to a conceptualized facies distribution that is anchored to the wells using a stratigraphic sequence approach. We have 100% cored porosity and permeability values in the COSTA model at the well level but, unfortunately, no access to the actual core for facies description. For this reason, the inter-well regions of the model used a generated and not conceptualized facies distribution to trend the properties, i.e. permeability. We developed pseudo facies distributions by modelling the mean PTSD (in microns) across the entire model using the Winland R35 equation. By generating mean PTSD maps, we could reverse engineer the pseudo facies distributions by discretising specific pore throat sizes with a lithological class, i.e. >7 microns would be grain supported high energy carbonate fabrics whilst <1 micron would be more mud-supported low energy carbonate fabrics. The facies distributions are referred to as pseudo because they were purely based on petrophysics to be able to reconcile drainage capillary pressure per RRT for similar depositional environments (low and high energy) and petrophysical trends affected by diagenetic overprints on the Upper Kharaib Member (Dunnington 1967; Harris et al. 1968; Oswald et al. 1995; Grotsc 1998; Davies et al. 2002; Strohmenger et al. 2004, 2006, 2007; Venkitadri et al. 2005; Saadawi 2006; Yose et al. 2006; Al Mansoori et al. 2008; Ehrenberg et al. 2008; Buchem et al. 2010a, b; Lawrence et al. 2010; Yin et al. 2010; Koppen et al. 2015; Paganoni et al. 2015; Wazir et al. 2015; Wu and Ehrenberg 2016; Jeong et al. 2017; Jingjing et al. 2018).

Recall that all 43 wells used in this study originate from several different fields producing across the shelf-to-basin profile of the Upper Kharaib Member. Due to the nature of our dataset, defining a single spatial experimental variogram for such a vast area, i.e. at the basin-scale, with many lateral geological variations, is challenging. For example, the experimental spatial variogram for porosity continuously increases with lag distance; however, the data starts to enter a new cycle at around 32 km in the major direction (Fig. 9). The trough in the data indicates a non-monotonic variogram structure otherwise known as a ‘hole effect’ structure (Journel and Huijbregts 1978; Pyrcz and Deutsch 2003). The origin of the hole...
effect structure in our dataset is that we enter wells from another field with a different trend after a certain distance. The more wells we include to determine the experimental variogram, i.e. 60 km, the less correlatable the data becomes. Therefore, this study focused on the first data samples (smaller distances c. 32 km major range), which are more correlatable and have similar trends.

In this study, each ensemble has six uniquely generated carbonate build-up scenarios (i.e. facies maps) for each of the 62 geological layers (Fig. 10). We developed them by first modelling cored porosity and permeability independently in 3D using the same spatial geostatistical models and parameters using Sequential Gaussian Simulation (SGS). We repeated this step using all vertical permeability profiles derived from different permeability well-log upscaling techniques (harmonic, geometric, and arithmetic). We then calculated the mean PTSD (in microns) using the Winland R35 equation in 3D. The entire process was repeated once again using wireline porosity data. The individual PTSD maps were then converted from continuous data into discrete data (RRTs) by discretising the entire PTSD histogram into 27-micron ranges (0–7 µm), resulting in 27 RRTs.

While all models originally had 27 RRTs, all ensemble members of the open-source package have reduced RRTs. We discretised the 27 RRTs into 3, 5, and 7 RRT models by screening the RRT histograms and grouping certain RRTs with one another in such a way as to preserve the large-scale heterogeneity (Fig. 10). Therefore, each ensemble member has a version with 3, 5, and 7 RRTs. The implication of having variable amounts of RRTs for a given model concept could directly influence the sweep and drainage during reservoir simulation and therefore helps expand the uncertainty envelope of this study.

We modelled uncertainties in relative permeability curves by reducing the reservoir rock type maps from 27 RRTs in the ‘truth case’ down to 7, 5, and 3. The ensemble members with 7, 5, and 3 RRTs mimic the average wettability trend along the reservoir but not the exact resolution as our ‘truth case’ with 27 RRTs. In other words, we did not create model uncertainties in relative permeabilities by changing endpoints or initial/residual saturations but through the number of relative permeability curves according to the RRTs.

Once we had our full suite of RRT maps, cross-plots of 27 PTSD ranges from Winland R35 were then used as poro-perm transforms. Each ensemble has nine permeability maps, six bivariate, and three non-bivariate distributions. The 3D permeability distribution was modelled using collocated co-kriging SGS for each vertical permeability profile (harmonic, geometric, and arithmetic), honouring the spatial geological trends of their respective RRT maps. The spatial bivariate permeability distributions were linked with either wireline or cored porosity values. We also included three non-bivariate permeability distributions per ensemble because permeability prediction at the well level was not an issue.

**Saturation height modelling**

The entire process of generating RRT maps and quantifying how many RRTs would eventually be used was synergetic with
screening/grouping the data from our synthetic SCAL database. Our database contains 110 capillary pressure and PTSD curves from anonymised MICP carbonate plug samples, all linked with porosity, permeability and Winland R35 values. Unimodal, bimodal and trimodal PTSDs are present, capturing multiple realistic drainage capillary pressure curves for a wide range of carbonate pore fabrics. The open-source provision of the underpinning raw data provides users with the opportunity to test the impact of using different saturation height modelling functions on STOOIP and production. For each of the 27 RRTs, a minimum of three capillary drainage curves were selected from our SCAL database based on similar Winland R35 values. There are two reasons for doing so. The first reason was diagenetic overprinting. We picked a better or poorer drainage capillary pressure curve for a given RRT to mimic dissolution (i.e. permeability enhancement) or cementation (i.e. permeability reduction). The second reason was to create some real-life uncertainty in selecting the best representation of the water saturation with height above FWL for a given RRT/grid cell. Over 53 trillion combinations of capillary pressure curves and RRTs are possible for model initialisation from which an undisclosed combination was selected for the STOOIP calculation of our ‘truth case’. The STOOIP of our undisclosed ‘truth case’ was generated from 27 saturation height functions modelled with the Skelt-Harrison approach (Harrison and Jing 2001). A GRG Non-linear Solver was used to back-calculate the most appropriate Skelt-Harrison fitting parameters for a given drainage capillary pressure curve. Although other models can be more suitable for capturing bimodal and trimodal pore throat size distributions, the saturation model and the capillary-driven reservoir rock typing methodology used in this study proved to be comparatively good at matching the saturation profiles from the open-hole logs. Comparative results are exemplified in the supplementary material. The range in STOOIP for all ensemble members with original 27 RRT distributions is ±16%. Each ensemble member, which has reduced, i.e. 3, 5, or 7 RRTs, was deliberately not matched with the open-hole logs and would need to be corrected during history matching. In this work, we also incorporated uncertainties in the model ensembles related to the number of OWCs and the thicknesses of the capillary transition zones. Recall that our undisclosed ‘truth case’ has 27 RRTs, and each ensemble member which is part of the released data has a reduced number of RRTs. A smaller number of RRTs signifies a reduced resolution of saturation distribution with height above the FWL. Therefore, the number of OWCs and the thickness of the capillary transition zones inevitably differ from our ‘truth case’.

**Dynamic reservoir models**

**Overview**

A commercial three-phase, black-oil reservoir simulator was used to perform flow simulations. All models were initialised as an undersaturated system with a gravity-capillary equilibrium and active hysteresis to simulate natural depletion and waterflooding schemes. Pressure support is provided by an analytical aquifer to add realism and complexity and was modelled using the Carter-Tracy method. Multiple oil production plateau scenarios were tested to quantify the ideal oil plateau rate, which turned out to be 280,000 bbl/day for an average of 12 years. Hence, 280,000 bbl/day was considered the maximum constraint of the surface facilities. Each producer was assumed to have a maximum oil rate of 5000 bbls/day and a BHP constraint of 300 psi above the bubble point. The water injectors were solely constrained to a BHP of 6000 psi. The complete set of reservoir model parameters and inputs are shown in the supplementary material.

**Dynamic reservoir model design**

We kept the 27 RRT distribution from our ‘truth case’ to generate a flow behaviour that is as realistic as possible. We created oil–water relative permeability and forced imbibition capillary pressure curves.
(hysteresis) for each RRT. The forced imbibition capillary pressure and the oil-water relative permeability curves are synthetic but attempt to mimic the complexity of carbonate rock wettability which can be oil-wet, neutral, or mixed-wet. We account for hysteresis by defining (forced) imbibition capillary pressure curves to create more realistic production profiles, including residual oil saturation values for each RRT after waterflood. These were created using a modified Skjaeveland equation (Skjaeveland et al. 2000).

We used the modified Brooks and Corey equations (Brooks and Corey 1964) to generate oil-water relative permeability curves. The relative permeability and capillary pressure curves were generated by honouring the capillary drainage curves used for model initialisation which matches the saturation profiles of the open-hole logs. Their respective irreducible water saturation cut-offs are honoured through the endpoint relative permeabilities to ensure that the data is geologically and petrophysically consistent. The complete set of input curves is shown in the supplementary material.

PVT data was established based on synthetic but realistic generalized values found from fields produced from the Upper Kharab Member. We used a light oil density of 50.86 lb ft$^3$ produced as an undersaturated system with aGOR of 731 scf/STB. Additional information for the PVT properties is provided in the supplementary material.

**Field development plan**

30 years of synthetic production data (from 2020–50) was generated for one of the three anticlines belonging to our undisclosed ‘truth case’. The FDP schematic is shown in Figure 11. Production comes solely from 17 appraisal wells under natural depletion during the first five years. A two-stage 15-year FDP (from 2025–40) added 5-spot water-flood patterns and peripheral water injectors to the model. The first stage of the FDP assumes two drilling rigs from 2025 to 2029. The second stage assumes a third drilling rig available from 2030 to 2040. On average, each drilling rig added ten vertical onshore wells per year (including mobilization) to a target depth of 7374 to 8310 ft TVDs to mimic realistic onshore drilling campaigns typical for giant onshore carbonate reservoirs in the Middle East.

430 vertical wells were added to the model over 15 years; 248 production wells and 182 water injectors. The water injectors are divided into 88 peripheral and 94 infill wells. The peripheral water injectors contour the FWL and are roughly spaced 2 km apart, whilst the producers and infill water-injectors are all in a 5-spot water-flood pattern with 1 km spacing. Both the producers and peripheral water injectors perforate all 62 layers of the reservoir. However, the infill water-injectors perforate only the bottom 32 layers of the reservoir to avoid early water breakthrough due to the permeability contrast between the upper and lower parts of the reservoir.

We note that this FDP is not the optimum solution to manage this reservoir. We targeted long production plateau time and low water-cut as our management decision. On purpose, we leave a lot of room for reservoir simulation engineers and end-users to unravel the reservoir’s potential, i.e. allow them to explore different reservoir management decisions to improve the FDP further.

**Fig. 12.** Comparison of field rates for 72 ensemble members and the ‘truth case’ model (red line). The first two rows also show close-ups highlighted by the red box. All models have the same FWL to have a more focused look at the dynamic performance of the different model scenarios. The amount of reservoir rock types in all ensemble members varies between 3, 5, and 7.
Field performance

This section illustrates the dynamic field and well performance behaviours of 84 out of our 144 ensemble members.

Dynamic behaviour for long variogram range

All 72 open-source ensemble member models, which used a long major (32 km) and minor (20 km) variogram range to model petrophysical property distributions, were simulated to quantify the uncertainty envelope in field and well rates (Figs 12 and 13).

The maximum variability in oil production rate in 2031, i.e. before plateau production is reached, is c. 25%, equating to c. 40 000 bbl/day. By 2050, the variability in relative cumulative oil production reaches c. 17% (relative) while it is c. 22% for the water-cut (absolute), and c. 9% for the recovery factor (absolute). Field rates vary significantly across all 72 ensemble members and provide a wide uncertainty envelope for later history matching exercises.

The recovery factors in Figure 12 are much lower in all ensemble members than our ‘truth case’ because the original distribution of RRTs has been clustered into 3, 5, and 7 RRTs. Therefore, the relative permeability and capillary pressure tables for reservoir simulation have also been reduced. As previously mentioned, the saturation distributions were deliberately not matched with the open-hole logs and would have to be corrected during history matching. Figure 14 demonstrates all ensemble members’ field performance with 27 RRTs. The production behaviour is still very heterogeneous, and only the recovery factors are more alike (±2.5%).

Field rates shown in Figure 13 are from 6 example wells located across multiple structural elevations. In all 72 ensemble members, the production behaviour shows diverse oil rates and water-cuts commonly observed in Middle Eastern carbonate reservoirs. Water production mainly becomes an issue in wells located near the flanks of the reservoir. However, the water breakthrough times are quite dispersed. Some wells belonging to a specific ensemble member exhibit negligibly small water-cuts whilst others start to water out simultaneously. Likewise, the time on the plateau can vary by ±4 years in crestal wells.

Dynamic behaviour for reduced variogram range

We conducted a sensitivity analysis on the horizontal and vertical variability of the petrophysical properties by altering the variogram ranges. We re-modelled all 72 original ensemble members with a reduced horizontal anisotropic range of porosity and permeability of 6 km. Figure 15 illustrates the consequence of reducing only the horizontal range on the field performance of 12 model variants. Figure 16 demonstrates the consequence of reducing both the horizontal and vertical variogram range on the field performance for one ensemble member.

The maximum variability in field production rates due to a reduction in horizontal anisotropic range is summarized as follows (readings taken in the year 2050); oil production rate c. 74% (relative), water-cut c. 22% (absolute), and recovery factor c. 3% (absolute) ignoring lowest RF model anomaly. The time on plateau varies by ±13 years, and some models do not reach the 280 000 bbl/day plateau.

In Figure 16, we have a more focused look at the implication of reducing the horizontal continuity of the petrophysical properties for a particular ensemble member. Using a reduced variogram range underestimates the overall model pore volume by 6%, equating to...
1.8 billion barrels. The daily oil production is 13% less (33 000 bbl/day) in the year 2031, the time on plateau drops by four years, and overall, the model produces 2% higher water-cut in the year 2050.

We also analysed the impact of reducing the vertical and horizontal variogram range for the same model. The following observations are noticed: time on the plateau is reduced by one year along with a 0.6% smaller pore volume. The impact of reducing the vertical variogram range is less substantial than reducing the horizontal variogram range because the vertical heterogeneity (flow units) has already been captured in the basic framework model.

**Effect of free water level**

Recall that all 144 ensemble members have been assigned a unique variation in FWL of ±30 ft. While the saturation distributions can represent (at large) the saturation derived from the open-hole logs even with a variation in the FWL of ±30 ft, the impact on the reservoir dynamics is significant (Fig. 17). The time on plateau varies by ±4 years and the water-cut by ±20%.

**Grid refinement**

Recall that all model releases have 1.3 million active grid blocks. The grid cells are 250 × 250 m long, and their thicknesses vary between 0.8 and 8 ft. Each of the 144 open-source models was run parallel on 24 CPUs and took between 8 to 24 hours to simulate 30 years of production history. Although a 1.3 million grid cell model is more likely to be used in a real-world environment to keep computational challenges at bay, we wanted to test the impact of grid refinement on the predicted reservoir performance. We refined...
the grid in x and y direction by a factor two and four while keeping layer thicknesses the same, resulting in models with 6.6 million active grid blocks (125 × 125 m aerial resolution) and 22.6 million active grid blocks (62.5 × 62.5 m aerial resolution). With 22.6 million active grid blocks, the model took 28 days to run on 24 CPUs. A significant variation in field performance was observed; time on plateau ±5.5 years, oil production rate ±35%, oil production cumulative ±9.3%, water-cut ±5% and finally recovery factor ±1.3%. Visual results are shown in Figures 18 and 19. The synthetic production data that we generated came from a grid-refined version of our undisclosed ‘truth case’, with 22.6 million active grid blocks.

Field optimization

In this study, all 144 open-source model releases incorporate our 15-year FDP, which has 5-spot waterflood patterns along with peripheral water injectors contouring the FWL. As noted above, our FDP is by no means the optimum solution to manage this reservoir and can likely be improved through better water management, well placement, perforations, etc. We present eight different well configuration examples and their impact on field performance (Fig. 20). These eight scenarios were tested on a small sector of an ensemble member with a dimension of 30 km² (24 × 20 grid cells).

Cases 1 to 7 honour the following producer well constraints: they operate to a minimum BHP of 2200 psi and a maximum of 5000 bbl/day and constrain the maximum field production rate to 40 000 bbl/day (Fig. 20). Case 8 is the only scenario where we incorporate two deviated producer wells individually constrained to a maximum daily production rate of 20 000 bbl/day. All water injectors were constrained to operate to a maximum injection pressure of 6000 psi. All horizontal producer wells have a length of 1750 m, whilst in case 8, the deviated producer wells are 3750 m long. All horizontal water injectors are 2750 m long. Both water...
injectors and producers are fully perforated, and all wells are spaced 1.5 km apart. The dynamic performance between all eight cases is shown in Figure 21 and Table 1. We did not investigate economic costs. We only analysed the variations in field performance, i.e. water breakthrough time, water-cut, recovery factor, and voidage replacement. To optimize an ensemble member in its entirety (c. 1600 km²) will pose a unique challenge because of the multiple reservoir architectures with diverse lateral and vertical geological complexities. Fast screening tools might be needed to enable such studies (Scheidt and Caers 2009; Mayner et al. 2015).

**Summary**

We discussed the design of the COSTA model, a new open-source reservoir case study that includes realistic geological, petrophysical, and geomodelling uncertainties typical of aggradational carbonate build-up reservoirs. The COSTA model accounts for the many uncertainties inherent to the entire process of characterizing and modelling heterogeneous carbonates, which, as demonstrated, impact reservoir volumetrics and reservoir performance. The COSTA model contains 144 pre-built reservoir models, each with 30 years of production history, and the underpinning data are released in open-source format.
Fig. 19. Field performance affected by grid refinement for one ensemble member. Original model grid dimension 250 × 250 m (blue), 125 × 125 m (red), and 62.5 × 62.5 m (green).

Fig. 20. Eight different well configurations tested on a small 30 km² sector model. Cross-sectional view of the model permeability with c. 300 mD streaks in the upper part of the model (red). Oil producers are coloured in grey and water injectors in pink.

Fig. 21. Field production behaviour of the eight different well configurations shown in Figure 20.
The COSTA model hence offers a unique dataset to support research and training in the fields of reservoir characterization, reservoir modelling, reservoir simulation, and uncertainty quantification for subsurface energy applications, from oil and gas to geothermal, CCS and beyond. Additionally, this is a good dataset for field development optimization and reservoir performance prediction to test different recovery processes. The value of this study is also to explore the impact of varying fluid flow physics on sweep and recovery across multiple reservoir architectures with diverse lateral and vertical rock complexities. The dataset available in this study also provides many opportunities for those wishing to rebuild their reservoir models when testing, for example, the impact of alternative geological concepts, geostatistical modelling techniques, reservoir rock typing methodologies, or saturation height modelling approaches on reservoir volumetrics and dynamic behaviours. No history matching work has been performed in this study. Still, the significant variations observed in the production behaviour across the 144 models provide a broad and challenging production envelop for history matching. We also reveal the effect of the grid refinement on the field performance of our undisclosed ‘truth case’ model.

A key limitation of the COSTA model is the absence of faults or fractures. We limited different geostatistical parameters for petrophysical modelling and only considered reducing the variability of the vario-nor fractures. We limited different geostatistical parameters for production envelop for history matching. We also reveal the effect of the grid refinement on the field performance of our undisclosed ‘truth case’ model. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability
All data generated or analyzed during this study are included in this published article (and its supplementary information files).

### References

Adayi, T., Awoloye, A., Gomes, J.S., Parra, H. and Hu, J. 2019. Large scale modeling and assessment of the feasibility of CO2 storage onshore Abu Dhabi. Energy, 185, 653–670, https://doi.org/10.1016/j.energy.2019.07.052

Ali Blooshi, M. 2018. Fracture-Related Diagenesis: A Record of Fluid Flow Through the Thamama Group, UAE. Ali, M., Aidarbayev, S., Searle, M. and Watts, A. 2017. Subsidence History and Seismic Stratigraphy of the Western Musandam Peninsula, Oman-United Arab Emirates Mountains

Al Maoonee, A.I., Strohmeier, C.J., Ghani, A.A.M.A. and Mitchell, J. 2008. High-resolution sequence-stratigraphic correlation and facies characterization of Upper Thamama (Lower Cretaceous) reservoir units and non-reservoir zones, Abu Dhabi, United Arab Emirates. Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, November 2008, SPE-117905-MS, https://doi.org/10.2118/117905-MS

Alsharhan, A.S. 1993. Asab Field – United Arab Emirates, Rub Al Khali Basin, Abu Dhabi. AAPG Special Volumes (TR: Structural Traps VIII), 28

Alsharhan, A.S. 2014. Petroleum systems in the Middle East. Geologic Society, London, Special Publications, 392, 361–408, https://doi.org/10.1144/SP392.19

Arnold, D., Demyanov, V., Tatum, D., Christie, M., Rojas, T., Geiger, S. and Corbett, P. 2013. Hierarchical benchmark case study for history matching, uncertainty quantification and reservoir characterisation. Computers & Geosciences, 50, 4–15, https://doi.org/10.1016/j.cageo.2012.09.011

Bentley, M. 2016. Modelling for comfort? Petroleum Geoscience, 22, 3–10, https://doi.org/10.1144/pegeo2014-089

Brooks, R.H. and Corey, A.T. 1966. Hydraulic Properties of Porous Media. Hydrology Papers, Colorado State University.

Buchem, F., Al-Husseini, M., Maurer, F. and Droste, H. 2010a. Barremian–Aptian Stratigraphy and Hydrocarbon Habitat of the Eastern Arabian Plate – Volume I. Gulf Petrolink, Bahrain

Buchem, F., Al-Husseini, M., Maurer, F. and Droste, H. 2010b. Barremian–Aptian Stratigraphy and Hydrocarbon Habitat of the Eastern Arabian Plate – Volume II. Gulf Petrolink, Bahrain

Carvalho, A., Dias, A. and Ribeiro, M. 2011. Oil recovery enhancement for Thamama B Lower in an Onshore Abu Dhabi Field. Using conceptual models to understand reservoir behaviour. SPE Reservoir Characterisation and Simulation Conference and Exhibition, October 2011, Abu Dhabi, UAE, SPE-148268-MS, https://doi.org/10.2118/148268-MS

Castro, S., Caers, J. and Mukerji, T. 2005. The Stanford vs Reservoir. Stanford Center for Reservoir Forecasting (SCRF), Annual Report, 153–154

Christie, M. and Blunt, M. 2001. Tenth SPE comparative solution project: a comparison of upscaling techniques. SPE Reservoir Simulation Symposium, Society of Petroleum Engineers, 4, 308–317, SPE-72469-PA, https://doi.org/10.2118/72469-PA

Correa, M., Holendorff, J., Gaspar, A.T.F.S. and Schiozer, D. 2015. UNISIM-II: benchmark case proposal included in this published article (and its supplementary information files). Computers & Geosciences, 50, 4–15, https://doi.org/10.1016/j.cageo.2012.09.011

EAGE, London, 1206, https://doi.org/10.3997/2214-4609.201901631

Funding J. Costa Gomes’s PhD research was funded by a James Watt Scholarship and Energi Simulation. S. Geiger’s chair was funded by Energi Simulation.

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
