A PRP-HS Type Hybrid Nonlinear Conjugate Gradient Method for Solving Unconstrained Optimization Problems

Olawale J. Adeleke[1(✉)], Micheal O. Olusanya[2], and Idowu A. Osinuga[3]

[1] Department of Mathematics, College of Science and Technology, Covenant University, Ota, Nigeria
wale.adeleke@covenantuniversity.edu.ng
[2] Department of Information Technology, Faculty of Accounting and Informatics, Durban University of Technology, Durban, Republic of South Africa
[3] Department of Mathematics, College of Physical Sciences, Federal University of Agriculture, Abeokuta, Nigeria

Abstract. Many engineering problems that occur in real-life are usually constrained by one or more factors which constitute the basis for the complexity of obtaining optimal solutions. While some of these problems may be transformed to the unconstrained forms, there is a large pool of purely unconstrained optimization problems in engineering which have practical applications in the industry. One effective approach for solving this latter category of problems is the nonlinear conjugate gradient method (NCGM). Particularly, the NCGM uses an efficient recursive scheme to solve unconstrained optimization problems with very large dimensions. In this paper, a new hybrid NCGM is proposed based on the recent modifications of the Polak-Ribiére-Polyak (PRP) and Hestenes-Stiefel (HS) methods. Theoretical analyses and numerical computations using standard benchmark functions, as well as comparison with existing NCGM schemes show that the proposed PRP-HS type hybrid scheme is globally convergent and computationally efficient.

Keywords: Hybrid methods · Nonlinear conjugate gradient method · Unconstrained optimization problems · Descent property · Global convergence · Line search

1 Introduction

This study considers the nonlinear conjugate gradient method (NCGM), an iterative method for solving the unconstrained optimization problem of the form

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1}$$

where the gradient of the continuous function $f : \mathbb{R}^n \to \mathbb{R}$ can be evaluated at every point The method and all its variants constitute a class of iterative scheme for solving

R. Silhavy et al. (Eds.): CoMeSySo 2019, AISC 1046, pp. 58–68, 2019.
https://doi.org/10.1007/978-3-030-30329-7_6

(1) effectively especially for large values of n. The NCGM scheme for (1) can be written as follows:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{2}$$

where the stepsize, α_k, evaluated by a suitable line search process, and the search direction, d_k, is computed by the formula

$$\begin{cases} d_0 = g_0, & \text{for } k = 0 \\ d_k = -g_k + \beta_k d_{k-1}, & \text{for } k \geq 1, \end{cases} \tag{3}$$

where g_k is the gradient of f at x_k, β_k is the NCGM update parameter and is usually chosen so that the (2) and (3) together reduces to the linear form of the method when f is a strictly convex quadratic function and α_k is determined by a one-dimensional exact line search technique.

Some of the variants of NCGM include, Hestenes and Stiefel (HS) [1], Fletcher and Reeves (FR) [2], Polak, Ribieré and Polyak (PRP) [3, 4], Liu and Storey (LS) [5], Dai and Yuan (DY) [6] and Hager and Zhang (CG_DESCENT (N)) [7]. These methods are computed as follows: $\beta_k^{HS} = g_k^T y_k / d_{k-1}^T y_{k-1}$, $\beta_k^{FR} = \|g_k\|^2 \Big/ \|g_{k-1}\|^2$, $\beta_k^{PRP} = g_k^T y_k \Big/ \|g_{k-1}\|^2$, $\beta_k^{LS} = -g_k^T y_k / d_{k-1}^T g_{k-1}$, $\beta_k^{DY} = \|g_k\|^2 \Big/ d_{k-1}^T y_{k-1}$ and $\beta_k^N = (y_k - 2d_{k-1} \left(\|y_k\|^2 \Big/ d_{k-1}^T y_k \right))^T g_k / d_{k-1}^T y_k$, where $\|\cdot\|$ denotes the Euclidean norm and $y_k = g_k - g_{k-1}$.

These variants are equivalent for strictly convex functions and exact line search. However, for non-quadratic functions, the methods perform differently [8, 9]. Establishing the global convergence results of these methods usually requires that the stepsize α_k satisfies certain approximate line search criteria as most exact line search procedures are cost-ineffective. The most widely used of these conditions are the strong Wolfe-Powell inequalities given as follows:

$$\begin{cases} f(x_{k-1}) - f(x_k) \geq -\delta \alpha g_{k-1}^T d_{k-1} \\ \left| g_k^T d_{k-1} \right| \leq -\sigma g_{k-1}^T d_{k-1}, \end{cases} \tag{4}$$

with $0 < \delta < \sigma < 1$. The weaker version of this condition is obtained by combining the first part of (4) with

$$g_k^T d_{k-1} \geq \sigma g_{k-1}^T d_{k-1}. \tag{5}$$

For the NCGM and its variant to be considered efficient, the search direction d_k, must satisfy a property known as the descent property. That is, d_k making an angle of strictly less than $\pi/2$ radians with $-g_k$, will always guaranteed a decrease in f provided that α_k is sufficiently small. In other words, an NCGM algorithm is said to satisfy the descent property if

$$g_k^T d_{k-1} < 0, \ \forall k \geq 1. \tag{6}$$

A more natural way of ensuring descent for NCGM algorithms is through the use of sufficient descent property. This is given by

$$g_k^T d_{k-1} \leq -c\|g_{k-1}\|^2, \ \forall k \geq 1, \tag{7}$$

where c is a positive constant.

Each CG method has very striking features that makes it adaptable to some sets of unconstrained problems. For instance, the FR and DY methods have been identified as having the best convergence results (See [10] and [11]). However, for general objective functions, the two methods have poor computation power. Conversely, the HS and PRP methods have good computational strength even though they exhibit poor convergence results. These contrasting features have led to the development of hybrid methods which are constructed with the aim of overcoming existing deficiencies in two or more methods. For instance, a well-constructed hybrid method of FR and PRP should perform well computationally as well as yield good convergence properties. This is the main motivation behind this paper. In this paper, a new hybrid NCGM variant is proposed through the combination of recently proposed variants of the traditional PRP and HS methods. Even though the two methods are specifically known to perform well computation-wise, it is interesting to know that the hybrid methods proposed here exhibit strong global convergence properties.

Recently, Du et al. [12] proposed four modified NCGMs. Only two which are of interest are mentioned in this study. The first of these, a variant of the PRP method which shall denoted as DPRP, has its β_k value given as follows:

$$\beta_k^{DPRP} = g_k^T \left(g_k - \left(\left| g_k^T g_{k-1} \right| \Big/ \|g_{k-1}\|^2 \right) g_{k-1} \right) \Big/ \|g_{k-1}\|^2 \tag{8}$$

and the second, denoted as DHS is a variant of the HS method and has its β_k given as follows:

$$\beta_k^{DHS} = g_k^T \left(g_k - \left(\left| g_k^T g_{k-1} \right| \Big/ \|g_{k-1}\|^2 \right) g_{k-1} \right) \Big/ d_{k-1}^T y_k. \tag{9}$$

It was further established in their study that the DPRP and DHS methods possess the sufficient descent property (Eq. 7) and are globally convergent under strong Wolfe line search with $0 < \sigma < 1/4$ and $0 < \sigma < 1/3$, respectively. It was also shown that the two methods are always nonnegative under the same assumptions of strong Wolfe line search and the respective intervals of convergence above.

Based on the DPRP and DHS methods, in this paper, a mixed hybrid NCGM is proposed. The development of this hybrid method was motivated by the observation that the numerical results reported in Du et al. [12] for the DPRP and DHS methods revealed that the methods are not so efficient compared to some classical methods, especially the CG_DESCENT method. The objective of this study is therefore, to construct a computationally efficient NCGM by hybridizing the PRP, HS, DPRP and

DHS methods. The remaining parts of this paper are organized as follow. The proposed hybrid method and the corresponding algorithm is presented in Sect. 2, while the descent and global convergence properties are described in Sect. 3. In Sect. 4, the numerical computations as well as the discussion of results are reported. The concluding remarks are given in Sect. 5.

2 Proposed Hybrid CG Methods and Algorithm

In this section, the new hybrid method of PRP, HS, DPRP and DHS is presented. A suitable algorithm for implementing the method for unconstrained optimization test problems is also provided in this section. Motivated by the ideas of hybrid method construction in [8, 13, 16] and the mixed performance profile of these methods, the β_k of the hybrid of DPRP, DHS, PRP and HS, denoted as HNCG, is given as

$$\beta_k^{HNCG} = \left(\|g_k\|^2 - \max\left\{\left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) g_k^T g_{k-1}, g_k^T g_{k-1}\right\}\right) \Big/ \max\left\{\|g_{k-1}\|^2, d_{k-1}^T (g_k - g_{k-1})\right\}. \tag{10}$$

By observation, the HNCG method represented in (10) can be reduced to any of PRP, DPRP, HS and DHS. A suitable algorithm for this method is given as follows:

Algorithm 1: Hybrid HNCG Scheme

1: input $x_0 \in \mathbb{R}^n$, $\varepsilon \geq 0$, set $d_0 = -g_0$, $k = 0$;
2: obtain α_k so that (4) is satisfied;
3: while $\|g_k\| < \varepsilon$ do
4: $y_k = g_k + g_{k-1}$;
5: $\beta_k = \beta_k^{HNCG}$ (as in (10));
6: $d_k = -g_k + \beta_k^{HNCG} d_{k-1}$;
7: $x_{k+1} = x_k + \alpha_k d_k$;
8: $k = k + 1$;
9: end

3 Descent and Global Convergence Properties of the HNCG Method

3.1 Descent Property of Algorithm 1

The descent property of d_k and the global convergence results for Algorithm 1 are presented in this section. The following result establishes the descent properties of the

sequence of search directions generated by the Hybrid Algorithm. Interestingly, this result was obtained independent of any line search procedure.

Theorem 1: Suppose that $f(x)$ in (1) is a smooth function and d_k is generated by the Hybrid Algorithm. Then $g_k^T d_k < 0$ for each $k \geq 0$.

Proof: Clearly, $g_0^T d_0 = -\|g_0\|^2 < 0$ for $k = 0$. By assumption, let $g_k^T d_k < 0$ holds for $k \geq 1$. To show that $g_k^T d_k < 0$ for all k, the proof is divided into four different cases for each of the hybrid method. Note that if $\beta_k = 0$, then, from (3), $d_k = -g_k + \beta_k d_{k-1} = -g_k \Rightarrow g_k^T d_k = -\|g_k\|^2 < 0$. Thus, it is assumed that $\beta_k \neq 0$ in all the cases.

Case I: If $g_k^T g_{k-1} < \left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) g_k^T g_{k-1}$ and $d_{k-1}^T (g_k - g_{k-1}) \geq \|g_{k-1}\|^2$, then from (10), $\beta_k^{HNCG} = \beta_k^{DHS}$. Notably, since $\|g_{k-1}\|^2 > 0$, the denominator of β_k^{DHS}, that is, $d_{k-1}^T (g_k - g_{k-1}) > 0$.. Hence, starting with (3), we have

$$\begin{aligned} g_k^T d_k = g_k^T \left(-g_k + \beta_k^{DHS} d_{k-1}\right) &= \left(\|g_k\|^2 g_{k-1}^T d_{k-1} \Big/ d_{k-1}^T (g_k - g_{k-1})\right) \\ &\quad - \left(\left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) \cdot g_k^T g_{k-1} \Big/ d_{k-1}^T (g_k - g_{k-1})\right) \cdot g_k^T d_{k-1} \\ &< \left(\|g_k\|^2 g_{k-1}^T d_{k-1} \Big/ d_{k-1}^T (g_k - g_{k-1})\right) < 0. \end{aligned} \tag{11}$$

In (11), the first equality was obtained by finding the inner product of (3) and substituting β_k^{DHS} for β_k. The first inequality is obvious because $\left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) > 1$ from the first assumption above, while the second was obtained from the fact that $g_{k-1}^T d_{k-1} < 0$.

Case II: If $g_k^T g_{k-1} < \left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) g_k^T g_{k-1}$ and $d_{k-1}^T (g_k - g_{k-1}) < \|g_{k-1}\|^2$, then (10) yields $\beta_k^{HNCG} = \beta_k^{DPRP}$. Notice that the second assumption implies that $g_k^T d_{k-1} - \|g_{k-1}\|^2 < g_{k-1}^T d_{k-1}$. Thus, starting from (3), we obtain the following result:

$$\begin{aligned} g_k^T d_k = g_k^T \left(-g_k + \beta_k^{DPRP} d_{k-1}\right) &= \left(\left(g_k^T d_{k-1} - \|g_{k-1}\|^2\right) \|g_k\|^2 \Big/ \|g_{k-1}\|^2\right) \\ &\quad - \left(\left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) \cdot g_k^T g_{k-1} \Big/ \|g_{k-1}\|^2\right) \cdot g_k^T d_{k-1} \\ &< \left(\left(g_k^T d_{k-1} - \|g_{k-1}\|^2\right) \|g_k\|^2 \Big/ \|g_{k-1}\|^2\right) < \left(\|g_k\|^2 \Big/ \|g_{k-1}\|^2\right) \cdot g_{k-1}^T d_{k-1} < 0. \end{aligned} \tag{12}$$

Case III: If $g_k^T g_{k-1} \geq \left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) g_k^T g_{k-1}$ and $d_{k-1}^T (g_k - g_{k-1}) \geq \|g_{k-1}\|^2$, then from (10), $\beta_k^{HNCG} = \beta_k^{HS}$. Proceeding from the inner product of (3) with g_k gives

$$\begin{aligned} g_k^T d_k = g_k^T\left(-g_k + \beta_k^{HS} d_{k-1}\right) = & \left(\|g_k\|^2 g_{k-1}^T d_{k-1} \Big/ d_{k-1}^T (g_k - g_{k-1})\right) \\ & - \left(g_k^T g_{k-1} \cdot g_k^T d_{k-1} \Big/ d_{k-1}^T (g_k - g_{k-1})\right) \\ & < \left(\|g_k\|^2 g_{k-1}^T d_{k-1} \Big/ d_{k-1}^T (g_k - g_{k-1})\right) < 0 \end{aligned} \tag{13}$$

Case IV: If $g_k^T g_{k-1} \geq \left(\left|g_k^T g_{k-1}\right| \Big/ \|g_{k-1}\|^2\right) g_k^T g_{k-1}$ and $d_{k-1}^T (g_k - g_{k-1}) < \|g_{k-1}\|^2$, then from (10), $\beta_k^{HNCG} = \beta_k^{PRP}$. The second assumption in this case allows us to set $g_k^T d_{k-1} - \|g_{k-1}\|^2 < g_{k-1}^T d_{k-1}$ and beginning from (3) as in other cases, we obtain

$$\begin{aligned} g_k^T d_k = g_k^T\left(-g_k + \beta_k^{PRP} d_{k-1}\right) & < \left(\|g_k\|^2 \left(g_k^T d_{k-1} - \|g_{k-1}\|^2\right) \Big/ \|g_{k-1}\|^2\right) \\ & < \left(\|g_k\|^2 \Big/ \|g_{k-1}\|^2\right) \cdot g_{k-1}^T d_{k-1} < 0 \end{aligned} \tag{14}$$

The descent property is satisfied in (11)–(14). Hence, the sequence of search directions generated by the HNCG method satisfies the descent property.

3.2 Global Convergence Property of Algorithm 1

The following important result is a consequence of Theorem 1 and is very important in establishing the global convergence of Algorithm 1. The result is stated as follows without proof.

Lemma 2: The inequality $0 \leq \beta_k \leq \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}$ always hold for every $k \geq 1$, where $\beta_k = \beta_k^{HNCG}$.

Next, the global convergence result for Algorithm 1 is presented. The result is established under the following assumptions.

Assumption:

1. Bound on the Objective Function: $f(x)$ is bounded from below on the level set $\Omega = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$, where x_0 is an initial guessed point.
2. Lipschitz Condition: within some neighbourhood N of Ω, $f(x)$ is continuously differentiable, and its gradient is Lipschitz continuous, that is, for all $x, y \in N$, there exists a constant $L \geq 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|. \tag{15}$$

3. Let an iterative scheme of the form (2) where d_k is a descent direction and α_k satisfies the Wolfe line search conditions (first part of 4) and (5). If Assumptions 1 and 2 hold, then,

$$\sum_{k=0}^{\infty}\left(\left(g_k^T d_k\right)^2 \Big/ \|d_k\|^2\right) < +\infty \tag{16}$$

Assumption 3 is the well-known Zoutendijk condition, the proof of which may be found in [11]. Using Lemma 2 and these assumptions, the following global convergence result of Algorithm 1 is presented.

Theorem 3: Let d_k be generated by the iterative rules (2)–(3) with β_k computed as in (10). If Assumptions (1)–(3) hold, then, $\lim_{k\to\infty} \inf\|g_k\| = 0$.

Proof: Suppose by contradiction that the conclusion does not hold, that is, $\lim_{k\to\infty} \inf\|g_k\| \neq 0$ However, since $\|g_k\| > 0$, there exists a constant $\varsigma > 0$ such that $\|g_k\| \geq \varsigma$, $\forall\, k$.

Starting with (3), the following holds,

$$\|d_k\|^2 = \beta_k^2\|d_{k-1}\|^2 - 2d_k^T g_k - \|g_k\|^2. \tag{17}$$

Dividing both sides of (17) by $\left(d_k^T g_k\right)^2$ and using Lemma 2, the following was obtained

$$\begin{aligned}\left(\|d_k\|^2 \Big/ \left(d_k^T g_k\right)^2\right) &= \left(\|d_{k-1}\|^2 \Big/ \left(d_{k-1}^T g_{k-1}\right)^2\right) - \left(2/d_k^T g_k\right) - \left(\|g_k\|^2 \Big/ \left(d_k^T g_k\right)^2\right) \\ &= \left(\|d_{k-1}\|^2 \Big/ \left(d_{k-1}^T g_{k-1}\right)^2\right) - \left((1/\|g_k\|) + \left(\|g_k\|/d_k^T g_k\right)\right)^2 + \left(1 \Big/ \|g_k\|^2\right) \\ &\leq \left(\|d_{k-1}\|^2 \Big/ \left(d_{k-1}^T g_{k-1}\right)^2\right) + \left(1 \Big/ \|g_k\|^2\right).\end{aligned} \tag{18}$$

Since the expression on the left side of (18) is the same as $1 \Big/ \left(d_1^T g_1\right)^2$, Eq. (18) gives

$$\left(\|d_{k-1}\|^2 \Big/ \left(d_{k-1}^T g_{k-1}\right)^2\right) \leq \sum_{i=1}^{k}\left(1 \Big/ \|g_i\|^2\right) \leq k/\varsigma,\ \forall\, k \tag{19}$$

Equation (19) implies that $\sum_{k\geq 1}\left(g_k^T d_k\right)^2 \Big/ \|d_k\|^2 = \infty$, which contradicts (16). □

4 Numerical Implementation of Algorithm 1

In this section, the numerical implementation of Algorithm 1 is presented. The indicators for comparing the algorithm with other existing methods are: the number of iterations (Table 1), the CPU time of executing the algorithm (Table 2), the optimal values of the objective function (Table 3) and the infinite norm of the gradient (Table 4). All the test functions were drawn from the Andrei [14], many of which are

also in the CUTEr library described in Bongartz et al. [15]. The algorithm was implemented on Matlab R2016a installed on a PC with Windows 10 OS and 2G RAM. For all the twenty selected problems, the dimension n is set to 5000 and 10000.

Table 1. Numerical test results indicating the number of iterations

Function name	Dim	DPRP	DHS	CGD	HNCG
EBD-1	5000	60	61	107	72
	10000	71	101	59	67
Diagonal-4	5000	91	8	52	18
	10000	91	29	53	18
EH	5000	140	95	33	29
	10000	105	77	33	30
GR	5000	39	56	57	64
	10000	39	56	57	64
EWH	5000	37	150	83	63
	10000	37	150	83	63
ET	5000	3	10862	180	628
	10000	3	13652	176	692
PQD	5000	3179	26	1590	37
	10000	965	25	483	36
LIARWHD	5000	18	21	30	21
	10000	18	22	31	23
QUARTC	5000	185	181	8	191
	10000	148	146	8	6
GWH	5000	37	150	84	63
	10000	37	150	88	63

The preliminary results show that the proposed hybrid method is efficient and competes very well with existing methods. For instance, by observing Table 1 where the numeral values of the numbers of iterations were reported, HNCG produces better results that the CG_DESCENT and DHS methods (HNCG vs. CG_DESCENT (14 to 6), HNCG vs. DPRP (8 to 12), HNCG vs. DHS (11 to 8, with one tie)). The same is true for the CPU time in Table 2 (HNCG vs. CG_DESCENT (15 to 5), HNCG vs. DPRP (10 to 10), HNCG vs. DHS (11 to 9)). The following acronyms were used in the tables: EBD = Extended Block Diagonal-1; EH = Extended Himmelblau; GR = Generalized Rosenbrock; EWH = Extended White & Holst; ET = Extended Tridiagonal; PQD = Perturbed Quadratic Diagonal; GWH = Generalized White & Holst; Dim = Problem dimension; CGD = CG_DESCENT.

Table 2. Numerical test results indicating the CPU time of Algorithm 1

Function name	Dim	DPRP	DHS	CGD	HNCG
EBD-1	5000	1.000	0.452	0.824	0.459
	10000	1.619	0.693	0.386	0.638
Diagonal-4	5000	1.530	0.112	0.492	0.161
	10000	2.212	0.348	0.606	0.207
EH	5000	2.215	0.890	0.264	0.242
	10000	2.468	1.137	0.405	0.375
GR	5000	0.429	0.540	0.651	0.611
	10000	0.686	0.865	1.086	0.972
EWH	5000	0.563	2.005	1.237	0.909
	10000	0.939	3.379	2.154	1.379
ET	5000	0.094	72.565	1.774	4.089
	10000	0.182	136.981	2.324	7.749
PQD	5000	24.268	0.113	6.145	0.129
	10000	16.431	0.216	4.162	0.280
LIARWHD	5000	0.125	0.131	0.212	0.340
	10000	0.475	0.318	0.463	0.343
QUARTC	5000	0.423	0.395	0.050	0.430
	10000	0.661	0.498	0.097	0.064
GWH	5000	0.359	1.145	0.660	0.484
	10000	0.649	2.153	1.305	0.875

Table 3. Numerical test results indicating the values of the objective function

Function name	Dim	DPRP	DHS	CGD	HNCG
EBD-1	5000	3.85e−15	6.79e−14	1.93e−16	4.03e−15
	10000	7.58e−15	2.57e−14	8.71e−14	2.46e−14
Diagonal-4	5000	5.58e−16	NaN	7.34e−16	5.89e−16
	10000	1.12e−15	5.10e−15	6.60e−16	1.18e−15
EH	5000	3.37e−15	3.29e−15	6.57e−15	4.24e−15
	10000	6.11e−15	1.19e−15	1.31e−14	7.34e−15
GR	5000	3.99e+00	3.99e+00	3.99e+00	3.99e+00
	10000	3.99e+00	3.99e+00	3.99e+00	3.99e+00
EWH	5000	1.24e−16	3.07e−16	5.99e−17	2.12e−16
	10000	1.24e−16	3.07e−16	5.99e−17	2.12e−16
ET	5000	0.00e+00	9.91e−09	6.58e−09	8.82e−09
	10000	0.00e+00	1.25e−08	4.34e−10	1.11e−08
PQD	5000	3.60e−16	6.43e−17	3.60e−16	1.67e−16
	10000	9.37e−17	4.98e−18	9.37e−17	6.00e−17
LIARWHD	5000	1.62e−23	1.20e−22	1.05e−22	4.83e−22
	10000	3.23e−23	0.00e+00	2.93e−23	2.32e−23
QUARTC	5000	9.13e−11	9.17e−11	6.61e−15	9.12e−11
	10000	7.28e−11	7.23e−11	1.32e−14	4.03e−12
GWH	5000	1.24e−16	3.07e−16	7.16e−17	2.12e−16
	10000	1.24e−16	3.07e−16	9.22e−17	2.12e−16

Table 4. Numerical test results indicating the values of norm of gradient

Function name	Dim	DPRP	DHS	CGD	HNCG
EBD-1	5000	1.97e−07	7.92e−07	8.15e−08	3.58e−07
	10000	4.93e−07	3.41e−07	8.26e−07	8.56e−07
Diagonal-4	5000	6.88e−07	NaN	7.93e−07	6.25e−07
	10000	9.73e−07	8.82e−07	7.52e−07	8.83e−07
EH	5000	8.43e−07	6.82e−07	6.84e−07	7.28e−07
	10000	9.50e−07	3.83e−07	9.67e−07	6.19e−07
GR	5000	7.29e−07	9.37e−07	8.22e−07	9.00e−07
	10000	7.29e−07	9.37e−07	8.22e−07	9.00e−07
EWH	5000	8.30e−07	9.85e−07	6.01e−07	8.97e−07
	10000	8.30e−07	9.85e−07	6.01e−07	8.97e−07
ET	5000	0.00e+00	9.96e−07	9.21e−07	9.92e−07
	10000	0.00e+00	9.98e−07	4.10e−07	9.94e−07
PQD	5000	9.84e−07	4.16e−07	9.84e−07	6.71e−07
	10000	9.85e−07	2.27e−07	9.85e−07	7.88e−07
LIARWHD	5000	1.62e−07	4.41e−07	4.13e−07	8.85e−07
	10000	4.58e−07	0.00e+00	4.36e−07	3.88e−07
QUARTC	5000	9.94e−07	9.97e−07	7.80e−10	9.92e−07
	10000	9.97e−07	9.92e−07	1.56e−09	1.14e−07
GWH	5000	8.30e−07	9.85e−07	6.63e−07	8.97e−07
	10000	8.30e−07	9.85e−07	7.16e−07	8.97e−07

5 Conclusion

This study proposed a new mixed hybrid conjugate gradient method for solving large-scale unconstrained optimization problems. The method was constructed using ideas from previously constructed hybrid method and taking as base methods two recently proposed methods of the Polak-Ribiére-Polyak (PRP) and Hestenes-Stiefel (HS) family of methods. Analyses revealed that the method satisfies the descent condition and is globally convergent. Numerical experiments showed that hybrid methods is computationally efficient compared to non-hybrid methods. As part of future work, the proposed method will be tested against existing hybrid methods using larger set of test functions.

References

1. Hestenes, M.R., Stiefel, E.L.: Method of conjugate gradients for solving linear systems. J. Res. Nat. Bur. Standards **49**, 409–436 (1952)
2. Fletcher, R., Reeves, C.: Function minimization by conjugate gradients. Comput. J. **7**, 149–154 (1964)
3. Polak, E., Ribieré, G.: Note sur la convergence de directions conjugeés. Rev. Francaise Informat Recherche Operationelle, 3e Année, vol. 16, pp. 35–43 (1969)

4. Polyak, B.T.: The conjugate gradient method in extreme problems. USSR Comput. Math. Math. Phys. **9**, 94–112 (1969)
5. Liu, Y., Storey, C.: Efficient generalized conjugate gradient algorithms, part 1: theory. J. Optim. Theory Appl. **69**, 129–137 (1991)
6. Dai, Y.H., Yuan, Y.: A nonlinear conjugate gradient method with a strong global convergence property. SIAM J. Optim. **10**, 177–182 (1999)
7. Hager, W.W., Zhang, H.: A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim. **16**(1), 170–192 (2005)
8. Adeleke, O.J., Osinuga, I.A.: A five-term hybrid conjugate gradient method with global convergence and descent properties for unconstrained optimization problems. Asian J. Sci. Res. **11**, 185–194 (2018)
9. Hager, W.W., Zhang, H.: A survey of nonlinear conjugate gradient methods. Pac. J. Optim. **2**, 35–58 (2006)
10. Al-Baali, M.: Descent property and global convergence of the Fletcher-Reeves method with inexact line search. IMA J. Numer. Anal. **5**(1), 121–124 (1985)
11. Dai, Y.H., Yuan, Y.: An efficient hybrid conjugate gradient method for unconstrained optimization. Ann. Oper. Res. **103**, 33–47 (2001)
12. Du, X., Zhang, P., Ma, W.: Some modified conjugate gradient methods for unconstrained optimization. J. Comput. Appl. Math. **305**, 92–114 (2016)
13. Jian, J., Han, L., Jiang, X.: A hybrid conjugate gradient method with descent property for unconstrained optimization. Appl. Math. Model. **39**, 1281–1290 (2015)
14. Andrei, N.: An unconstrained optimization test functions collection. Adv. Mod. and Optim. **10**(1), 147–161 (2008)
15. Bongartz, I., Conn, A.R., Gould, N.I.M., Toint, P.L.: CUTE: constrained and unconstrained testing environments. ACM Trans. Math. Software **21**, 123–160 (1995)
16. Oladepo, D.A., Adeleke, O.J., Ako, C.T.: A mixed hybrid conjugate gradient method for unconstrained engineering optimization problems. In: Silhavy, R. (eds.) Cybernetics and Algorithms in Intelligent Systems, CSOC2018 2018. AISC, vol. 765. Springer, Cham (2019)

Smart Injector Well Optimization for a Non-communicating Reservoir

D. A. Oladepo[1(✉)], C. T. Ako[1], O. D. Orodu[1], O. J. Adeleke[2], and A. S. Fadairo[1]

[1] Department of Petroleum Engineering, College of Engineering, Covenant University, Ota, Nigeria
david.oladepo@covenantuniversity.edu.ng
[2] Department of Mathematics, College of Science and Technology, Covenant University, Ota, Nigeria

Abstract. This study proposes a technique called intelligent well completions which uses water injection to improve sweep efficiency and recovery factor in oil production. The application of the technique in a water-flooding operation aims to optimize the outflow control valve (OCV) settings and injection rate of each segment of the oil production reservoir. A dynamic reservoir model was built using a reservoir flow performance simulator. It was observed that smart injection wells yield a better sweep efficiency with a favourable mobility ratio and improved pressure maintenance leading to an increase in field oil efficiency (FOE) when compared to the conventional wells and overall increase in net present value (NPV). In addition, the reservoir field oil efficiency was increased by 5% using the proposed technique.

Keywords: Field oil efficiency · Smart wells · Outflow Control Valve · Net present value

1 Introduction

Smart wells are non-conventional wells that make use of downhole sensors to monitor downhole temperature, pressure, flow saturation changes, phase composition. They are also completed with subsurface inflow control valves (ICVs) to control the flow of fluid from a segment of the reservoir to the well. They give real time information about the pressure, temperature, flow and phase composition at the subsurface (Fombad 2016). The smart well was originally used for shut off areas that were watered-out and as a measure taken against the early breakthrough during production. Improvements in smart well technologies and increasing amount of opportunities in more challenging assets, and the use of smart well technologies to improve recovery has caught significant attention in the oil and gas industry in the last decade. Several workflows have been developed and proposed in order to automate the whole process that integrates several sub processes focusing on specific parts of the surface or subsurface phenomena. Reservoir sweep is a crucial part of recovery efficiency, especially where significant investment is done by means of installing smart wells that feature ICVs,

R. Silhavy et al. (Eds.): CoMeSySo 2019, AISC 1046, pp. 226–232, 2019.
https://doi.org/10.1007/978-3-030-30329-7_21

which are remotely controllable. However, as it is a relatively newer concept, effective use of this technology has been a challenge (Ranjith et al. 2017).

Brouwer (2001) conducted a study to increase recovery through water flooding with smart well technology. In his work, water flooding was used to expand oil recovery after principal exhaustion. Zones with excessive permeability can affect the recovery of hydrocarbon, since they can bring about early water breakthrough and catching of by-passed oil. Smart well innovation provides the chance to balance these impacts by forcing a suitable pressure along the injection and production wells. In his research, water flooding is enhanced by changing the well profiles as indicated by some basic calculations that move stream ways far from the high porousness zone with a specific end goal to postpone water achievement. Aitokhuehi (2004) presented a study on the real-time optimization of smart by considering the use of both valve optimization and history matching. The author claimed that maximizing recovery in smart wells requires the estimation of the optimum settings of the control valves.

In another study, Almeida et al. (2007) used evolutionary optimization to control smart wells under technical uncertainties. Their work focused on an evolutionary algorithm that has the ability to optimize the mechanism for controlling the smart well technology currently used on the smart fields. El-Sayed et al. (2014) presented a study on the first intelligent well installation in UAE offshore field. The well was equipped with three inflow control valves and four permanent downhole gauges with remote control ability to control the flow from each zone monitoring the real time gauge data. The completion enabled commingled production from three reservoirs while balancing flow contribution from each reservoir and avoiding cross-flows from one reservoir into another. The intelligent well technology is one of the most important technologies with the capability of optimizing reservoir management, increasing recoverable reserve, enhancing oil recovery and speeding up the oil exploitation.

Goodyear et al. (1996) investigated on hot water flooding for high permeability viscous oil fields. In their investigated work, they discussed about reservoirs at the North Sea having oil of high viscosity present in permeable sands, they then stated that the technique they feel best to be recommended is injecting hot water along with the application of Horizontal wells of high length. Christensen et al. (2000) investigated the compositional simulation of water-alternating-gas processes. Their study incorporated the "compositional simulator" and "calculation process" to establish the relative permeability for the injection of WAG. The performance of the Hysteresis is executed for the water and gas stages (wetting and non-wetting stage). Oladepo et al. (2017) presented a study that shows how various WAG ratios are being varied with WAG injection program. From their study, various WAG ratios were used in optimizing the ultimate recovery and were then compared with two secondary injection schemes. Ekebafe (2012) presented a studied on smart well technology application in deep-water field development. The study puts more emphasis on integrated method and the implementations that aided it. Their study provided further insights into field developments.

The remaining parts of this paper are organized as follows. In Sect. 2, the research methodology is presented for the problem under study. The various results obtained through modeling and simulations are presented and discussed in Sect. 3. The concluding remarks are given in Sect. 4.

2 Methodology

2.1 Dynamic Model Description

A dynamic model was built using Eclipse E-100 black oil simulator. The history matching, predictions and sensitivity analysis were carried out. The individual injection rates for each segment from the intelligent injection well will be optimized for each of the layers to be considered in the reservoir by making use of the Frontsim application available in Eclipse. For this study, the optimization process will involve optimizing well allocation rates, OCV's, well location as control variables to improve the volumetric sweep efficiency of the Water flooding injection profile.

2.2 Numerical Reservoir Simulator

Schlumberger GeoQuest's Eclipse was used as the numerical reservoir simulator in this research. Eclipse can easily implement many types of production and economic constraints through its existing keywords. These constraints include economic production limit, production rate limit, and bottom hole pressure limit (BHP).

2.3 Optimizer and Links to Simulator

The optimizer, as indicated above, drives ECLIPSE for objective function evaluations. An interface establishes communication between the optimization routines and the simulator. Although the cases presented here are based on the maximization of the recovery factor and cumulative oil production, different objective functions, such as the net present value of the project, the minimization of water cut or the gas-oil ratio of individual wells, well groups or the entire field can easily be implemented. Multiple valves installed on different wells and their associated production rates/bottomhole pressures or injection rates/injection pressures can also be optimized along with the valve settings.

3 Results and Discussion

3.1 Case 1: FIELD X (Conventional Waterflood (Non-communicating))

In this case, the conventional water flood was performed, without downhole outflow control on injection. This case actuated on a field found to be towards total marginality due to rapid decline in pressure. There are three 40 ft thick producing units which are separated by 10 ft thick shale barriers. The reservoir dimensions are 2500 $\times$ 2500 $\times$ 50 ft3. The porosity across each of the fourteen layers was assumed to be

homogenous. The layers are of inhomogeneous permeability from top to bottom. The shale barriers are impermeable. The ratio of the vertical to horizontal permeability in each layer was taken to be 0.1. The displacement is unfavourable, with an endpoint mobility ratio of 3.5. Water was injected at a bottomhole pressure target of 10000 psi and subject to the maximum injection rate of 6 MSTB/d. Production was specified to occur at a target bottomhole pressure of 1000 psi, subject to a maximum oil rate of 8 MSTB/d. In both the controlled and uncontrolled cases, we reduced the current liquid production rate by 60% whenever a sharp increase in water cut is noticed. A minimum oil production rate of 400 STB/d was also imposed as an economic constraint. The simulations were performed for 14620 days (40 years).

3.2 Case 2: FIELD X (Smart Waterflood (CHOKE CASE))

In this case, water was injected into the reservoir using downhole outflow control valve for optimal performance of the injector well. Three valves were used for the three reservoirs in the field. Sensitivity analysis was done on the Outflow Control Valves (OCV's) to achieve the optimum valve setting with a total of 30 runs. Water was injected at a bottomhole pressure target of 10000 psi and subject to the maximum injection rate of 6 MSTB/d. Production was specified to occur at a target bottomhole pressure of 1000 psi, subject to a maximum oil rate of 8 MSTB/d. In both the controlled and uncontrolled cases, the current liquid production rate was reduced by 60% whenever a sharp increase in water cut is noticed. A minimum oil production rate of 400 STB/d was also imposed as an economic constraint. The simulations were performed for 14620 days (40 years).

The results obtained from implementing the methods described above are presented in form of tables and figures. Each of the results is discussed appropriately. Note that COP denotes cumulative oil production and FOE stands for field oil efficiency.

From Table 1, Case 2 (choke case) gave a Field Oil efficiency of 36% compared the conventional case and base case which gave a FOE of 32% and 12.5% respectively. An increase in FOE for the choke case is as a result of the optimal settings of the OCV for the choke case.

Table 1. Field X performance overview for a non-communicating reservoir

	COP (MMSTB)	FOE (%)
Base case	18.75	12.5
Case 1	48	32
Case 2	54	36

In addition, Fig. 1 shows water production due to smart completion in comparison to the conventional waterflood process. An overall increase in oil recovery due to optimization was also observed.

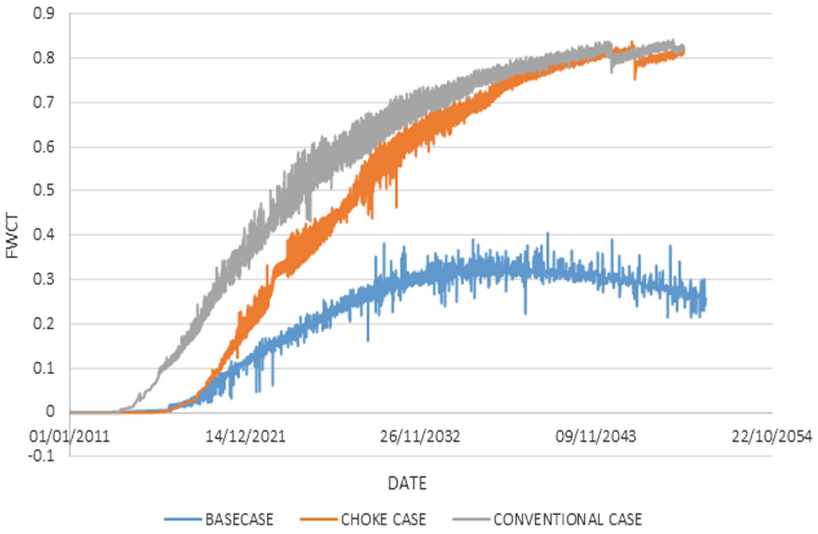

Fig. 1. Field water cut comparison case

3.3 Sensitivity of Valve Coefficient on Field Oil Efficiency Reservoir (Non Communicating)

This section describes the various results obtained from the sensitivity analysis performed on the valve coefficient. The analysis was conducted for the non-communicating well. VOC denotes valve opening coefficient.

Table 2 shows the sensitivity analysis done on different valve openings. Valve opening 0.3 gave us the optimal control setting for segment 1 with a corresponding FOE of 6.70% while at valve opening at 1 (fully opened valves) is at 6.50%.

Table 2. Segment 1 OCV size sensitivity analysis

VOC	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
FOE (%)	6.30	6.42	6.70	6.69	6.39	6.40	6.10	6.00	6.20	6.50

Table 3 shows the sensitivity analysis done for the OCV settings of Segment 2. It will be noticed that 0.5 optimal setting gave the highest FOE value and cumulative recovery compared with the OCV setting of 1 (fully opened), which gave an FOE of 16.2%.

Table 3. Segment 2 OCV settings sensitivity analysis

VOC	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
FOE (%)	16.50	16.42	16.70	16.69	16.40	16.10	16.10	16.00	16.40	16.20

3.4 Economic Results

From Fig. 2, it is easy to notice the improved economics of the field, a base case experiencing an increase in overall in oil revenue of about 140% for conventional waterflooding and 166.7% for smart development i.e. a net change of 26.7% from conventional waterflood to smart waterflood injection strategy, generally for non-communicating reservoir. Overall development with smart injection strategy yields best economic estimates as judged by the payout time (POT).

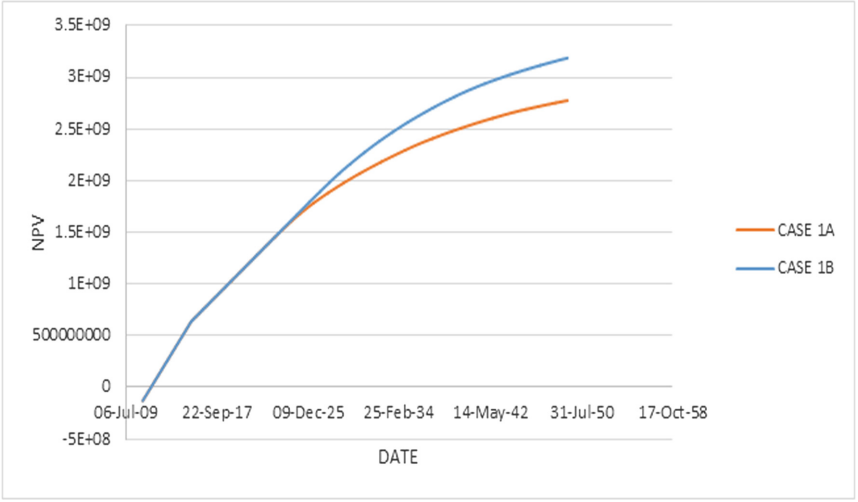

Fig. 2. Pay Out Time

It can be deduced that the POT for all case scenarios is averagely the same, but already established is the drawback of using the POT as an economic indicator, i.e. it does not take cognisance of the future project performance, hence not clearly representative, nonetheless smart well development strategy can be appreciated by virtue of the distinct differences in project performance. Also from the above indicator, it will be noticed that the performance of the NPV outputs increase with smart injection strategy.

4 Conclusion

The use of smart injector well in some oil producing regions of the world is relatively new and not much is known about its applications in terms of the control settings. Smart well technology aids better reservoir management and proper dynamics synthesis. This study has established a foundation for smart injector wells in terms of having an optimal setting for the outflow control valves. The major contributions of this project can be summarised as follows. The reactive control strategy used in this study for smart intelligent well resulted in a significant increase in oil recovery factor over the conventional well and also a cumulative improvement in production. Good sweep

efficiency was also achieved in each of the segments and also shows a significant increase in the total field oil production as well as an improved field water cut curve.

Acknowledgement. The authors would like to appreciate the support of Covenant University Management and Covenant University Centre for Research, and Innovation (CUCRID) for their support and funding the publication of this research output.

References

Almeida, L.F., Tupac, Y.J., Pacheco, M.A., Vellasco, M.M., Lazo, J.G.: Evolutionary optimization of smart-wells control under technical uncertainties. In: Latin American & Caribbean Petroleum Engineering Conference, 1 January 2007. Society of Petroleum Engineers (2007)

Fombad, M.W.: A technology perspective and optimized workflow to intelligent well applications. Ph.D. dissertation (2016)

Brouwer, D.R., Jansen, J.D., Van der Starre, S., Van Kruijsdijk, C.P., Berentsen, C.W.: Recovery increase through water flooding with smart well technology. In: SPE European Formation Damage Conference, 1 January 2001. Society of Petroleum Engineers (2001)

Christensen, J.R., Larsen, M., Nicolaisen, H.: Compositional simulation of water-alternating-gas processes. In: SPE Annual Technical Conference and Exhibition, 1 January 2000. Society of Petroleum Engineers (2000)

Ekebafe, A., Ogan, A.: Smart well technology application in deepwater field development. In: Nigeria Annual International Conference and Exhibition, 1 January 2012. Society of Petroleum Engineers (2012)

El-Sayed, M., Al Mutairi, A.M., Hassane, M.A., Kutty, S.M., Karrani, S.M., Kurian, A.: Three-zone commingled and controlled production using intelligent well completion. In: Abu Dhabi International Petroleum Exhibition and Conference, 10 November 2014. Society of Petroleum Engineers (2014)

Goodyear, S.G., Reynolds, C.B., Townsley, P.H., Woods, C.L.: Hot water flooding for high permeability viscous oil fields. In: SPE/DOE Improved Oil Recovery Symposium, 1 January 1996. Society of Petroleum Engineers (1996)

Aitokhuehi, I.: Real-time optimization of smart wells. Master of Science dissertation. Stanford university, June 2004

Oladepo, D.A., Churchill, A., Fadairo, A., Ogunkunle, T.F.: Evaluation of different wag optimization and secondary recovery techniques in a stratified reservoir. Int. J. Appl. Eng. Res. **12**(20), 9259–9270 (2017)

Ranjith, R., Suhag, A., Balaji, K., Putra, D., Dhannoon, D., Saracoglu, O., Hendroyono, A., Temizel, C., Aminzadeh, F.: Production optimization through utilization of smart wells in intelligent fields. In: SPE Western Regional Meeting, 23 April 2017. Society of Petroleum Engineers (2017)