Computational Materials Science 108 (2015) 233–238

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Editor's Choice

The AFLOW standard for high-throughput materials science calculations

Camilo E. Calderon [a], Jose J. Plata [a], Cormac Toher [a], Corey Oses [a], Ohad Levy [a,1], Marco Fornari [b], Amir Natan [c], Michael J. Mehl [d], Gus Hart [e], Marco Buongiorno Nardelli [f], Stefano Curtarolo [g,*]

[a] Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA
[b] Department of Physics, Central Michigan University, Mount Pleasant, MI 48858, USA
[c] Department of Physical Electronics, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel
[d] Center for Computational Materials Science, Naval Research Laboratory, Washington, DC 20375-5345, USA
[e] Department of Physics and Astronomy, Brigham Young University, Provo, UT 84602, USA
[f] Department of Physics and Department of Chemistry, University of North Texas, Denton, TX 76203, USA
[g] Materials Science, Electrical Engineering, Physics and Chemistry, Duke University, Durham, NC 27708, USA

ARTICLE INFO

Article history:
Received 31 May 2015
Received in revised form 6 July 2015
Accepted 7 July 2015

Keywords:
High-throughput
Materials genomics
AFLOWLIB
VASP

ABSTRACT

The Automatic-Flow (AFLOW) standard for the high-throughput construction of materials science electronic structure databases is described. Electronic structure calculations of solid state materials depend on a large number of parameters which must be understood by researchers, and must be reported by originators to ensure reproducibility and enable collaborative database expansion. We therefore describe standard parameter values for *k*-point grid density, basis set plane wave kinetic energy cut-off, exchange–correlation functionals, pseudopotentials, DFT+U parameters, and convergence criteria used in AFLOW calculations.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The emergence of computational materials science over the last two decades has been inextricably linked to the development of complex quantum–mechanical codes that enable accurate evaluation of the electronic and thermodynamic properties of a wide range of materials. The continued advancement of this field entails the construction of large open databases of materials properties that can be easily reproduced and extended. One obstacle to the reproducibility of the data is the unavoidable complexity of the codes used to obtain it. Published data usually includes basic information about the underlying calculations that allows rough reproduction. However, exact duplication depends on many details, that are seldom reported, and is therefore difficult to achieve.

These difficulties might limit the utility of the databases currently being created by high-throughput frameworks, such as AFLOW [1–3] and the Materials Project [4,5]. For maximal impact, the data stored in these repositories must be generated and represented in a consistent and robust manner, and shared through standardized calculation and communication protocols. Following these guidelines would promote optimal use of the results generated by the entire community.

The AFLOW (Automatic FLOW) code is a framework for high-throughput computational materials discovery [1–3,6], using separate DFT packages to calculate electronic structure and optimize the atomic geometry. The AFLOW framework works with the VASP [7–10] DFT package, and integration with the Quantum ESPRESSO software [11] is currently in progress. The AFLOW framework includes preprocessing functions for generating input files for the DFT package; obtaining the initial geometric structures by extracting the relevant data from crystallographic information files or by generating them using inbuilt prototype databases, and then transforming them into standard forms which are easiest to calculate. It then runs and monitors the DFT calculations automatically, detecting and responding to calculation failures, whether they are due to insufficient hardware resources or to runtime errors of the DFT calculation itself. Finally, AFLOW contains postprocessing routines to extract specific properties from the results of one or more of the DFT calculations, such as the band structure or thermal properties [12].

The AFLOWLIB repository [2,3,6] was built according to these principles of consistency and reproducibility, and the data it contains can be easily accessed through a REpresentational State Transfer–Application Programming Interface (REST–API) [3]. In this

* Corresponding author.
E-mail address: stefano@duke.edu (S. Curtarolo).
[1] On leave from the Physics Department, NRCN, Israel.

http://dx.doi.org/10.1016/j.commatsci.2015.07.019
0927-0256/© 2015 Elsevier B.V. All rights reserved.

paper we present a detailed description of the AFLOW standard for high-throughput (HT) materials science calculations by which the data in this repository was created.

2. AFLOW calculation types

The AFLOWLIB consortium [2] repository is divided into databases containing calculated properties of over 625,000 materials: the Binary Alloy Project, the Electronic Structure database, the Heusler database, and the Elements database. These are freely accessible online via the AFLOWLIB website [6], as well as through the API [3]. The Electronic Structure database consists of entries found in the Inorganic Crystal Structures Database, ICSD [13,14], and will thus be referred to as "ICSD" throughout this publication. The Heusler database consists of ternary compounds, primarily based on the Heusler structure but with other structure types now being added.

The high-throughput construction of these materials databases relies on a pre-defined set of standard *calculation types*. These are designed to accommodate the interest in various properties of a given material (e.g. the ground state ionic configuration, thermodynamic quantities, electronic and magnetic properties), the program flow of the HT framework that envelopes the DFT portions of the calculations, as well as the practical need for computational robustness. The AFLOW standard thus deals with the parameters involved in the following calculation types:

i. `RELAX`. Geometry optimizations using algorithms implemented within the DFT package. This calculation type is concerned with obtaining the ionic configuration and cell shape and volume that correspond to a minimum in the total energy. It consists of two sequential relaxation steps. The starting point for the first step, `RELAX1`, can be an entry taken from an external source, such as a library of alloy prototypes [15,16], the ICSD database, or the Pauling File [17]. These initial entries are preprocessed by AFLOW, and cast into a unit cell that is most convenient for calculation, usually the standard primitive cell, in the format appropriate for the DFT package in use. The second step, `RELAX2`, uses the final ionic positions from the first step as its starting point, and serves as a type of annealing step. This is used for jumping out of possible local minima resulting from wavefunction artifacts.

ii. `STATIC`. A single-point energy calculation. The starting point is the set of final ionic positions, as produced by the `RELAX2` step. The outcome of this calculation is used in the determination of most of the thermodynamic and electronic properties included in the various AFLOW databases. It therefore applies a more demanding set of parameters than those used on the `RELAX` set of runs.

iii. `BANDS`. Electronic band structure generation. The converged `STATIC` charge density and ionic positions are used as the starting points, and the wavefunctions are reoptimized along standardized high symmetry lines connecting special k-points in the irreducible Brillouin zone (IBZ) [18].

These calculation types are performed in the order shown above (i.e. `RELAX1` → `RELAX2` → `STATIC` → `BANDS`) on all materials found in the Elements, ICSD, and Heusler databases. Those found in the Binary Alloy database contain data produced only by the two `RELAX` calculations. Sets of these calculation types can be combined to describe more complex phenomena than can be obtained from a single calculation. For example, sets of `RELAX` and `STATIC` calculations for different cell volumes and/or atomic configurations are used to calculate thermal and mechanical properties by the Automatic Gibbs Library (AGL) [12], and the Automatic Phonon Library (APL) [1], which are methods that have been implemented within the AFLOW framework. In the following, we describe the parameter sets used to address the particular challenges of the calculations included in each AFLOW repository.

3. The AFLOW standard parameter set

The standard parameters described in this work are classified according to the wide variety of tasks that a typical solid state DFT calculation involves: Brillouin zone sampling, Fourier transform meshes, basis sets, potentials, self-interaction error (SIE) corrections, electron spin, algorithms guiding SCF convergence and ionic relaxation, and output options.

Due to the intrinsic complexity of the DFT codes it is impractical to specify the full set of DFT calculation parameters within an HT framework. Therefore, the AFLOW standard adopts many, but not all, of the internal defaults set by the DFT software package. This is most notable in the description of the Fourier transform meshes, which rely on a discretization scheme that depends on the applied basis and crystal geometry for its specification. Those internal default settings are cast aside when error corrections of failed DFT runs, an integral part of AFLOW's functionality, take place. The settings described in this work are nevertheless prescribed as fully as is practicable, in the interest of providing as much information as possible to anyone interested in reproducing or building on our results.

3.1. k-point sampling

Two approaches are used when sampling the IBZ: the first consists of uniformly distributing a large number of k-points in the IBZ, while the second relies on the construction of paths connecting high symmetry (special) k-points in the IBZ. Within AFLOW, the second sampling method corresponds to the `BANDS` calculation type, whereas the other calculation types (non-`BANDS`) are performed using the first sampling method.

Sampling in non-`BANDS` calculations is obtained by defining and setting N_{KPPRA}, the number of k-points per reciprocal atom. This quantity determines the total number of k-points in the IBZ, taking into account the k-points density along each reciprocal lattice vector as well as the number of atoms in the simulation cell, via the relation:

$$N_{KPPRA} \leqslant \min\left[\prod_{i=1}^{3} N_i\right] \times N_a \tag{1}$$

N_a is the number of atoms in the cell, and the N_i factors correspond to the number of sampling points along each reciprocal lattice vector, $\vec{b}_i$, respectively. These factors define the grid resolution, $\delta k_i \|\vec{b}_i\|/N_i$, which is made as uniform as possible under the constraint of Eq. (1). The k-point meshes are then generated within the Monkhorst–Pack scheme [19], unless the material belongs to the hP, or hR Bravais lattices, in which case the hexagonal symmetry is preserved by centering the mesh at the Γ-point.

Default N_{KPPRA} values depend on the calculation type and the database. The N_{KPPRA} values used for the entries in the Elements database are material specific and set manually due to convergence of the total energy calculation. The defaults applied to the `RELAX` and `STATIC` calculations are summarized in Table 1. These defaults ensure proper convergence of the calculations. They may be too stringent for some cases but enable reliable application within the HT framework, thus presenting a practicable balance between accuracy and calculation cost.

Table 1
Default N_{KPPRA} values used in non-BANDS calculations.

Database	STATIC	RELAX
Binary Alloy	N.A.	6000
Heusler	10,000	6000
ICSD	10,000	8000

For BANDS calculations AFLOW generates Brillouin zone integration paths in the manner described in a previous publication [18]. The k-point sampling density is the *line density* of k-points along each of the straight-line segments of the path in the IBZ. The default setting of AFLOW is 128 k-points along each segment connecting high-symmetry k-points in the IBZ for single element structures, and 20 k-points for compounds.

The occupancies at the Fermi edge in all non-RELAX type runs are handled via the tetrahedron method with Blöchl corrections [20]. This involves the N_{KPPRA} parameter, as described above. In RELAX type calculations, where the determination of accurate forces is important, some type of smearing must be performed. In cases where the material is assumed to be a metal, the Methfessel–Paxton approach [21] is adopted, with a smearing width of 0.10 eV. Gaussian smearing is used in all other types of materials, with a smearing width of 0.05 eV.

3.2. Potentials and basis set

The interactions involving the valence electron shells are handled with the potentials provided with the DFT software package. In VASP, these include Ultra-Soft Pseudopotentials (USPP) [22,23] and Projector-Augmented Wavefunction (PAW) potentials [24,25], which are constructed according to the Local Density Approximation (LDA) [26,27], and the Generalized Gradient Approximation (GGA) PW91 [28,29] and PBE [30,31] exchange–correlation (XC) functionals. The ICSD, Binary Alloy and Heusler databases built according to the AFLOW standard use the PBE functional combined with the PAW potential as the default. The PBE functional is among the best studied GGA functionals used in crystalline systems, while the PAW potentials are preferred due to their advantages over the USPP methodology. Nevertheless, defaults have been defined for a number of potential/XC functional combinations, and in the case of the Elements database, results are available for LDA, GGA-PW91 and GGA-PBE functionals with both USPP and PAW potentials. Additionally, there are a small number of entries in the ICSD and Binary Alloy databases (less than 1% of the total) which have been calculated with the GGA-PW91 functional using either the USPP or PAW potential. The exact combination of exchange–correlation functional and potential used for a specific entry in the AFLOWLIB database can always be determined by querying the keyword dft_type using the AFLOWLIB REST-API [3].

DFT packages often provide more than one potential of each type per element. The AFLOW standardized lists of PAW and USPP potentials are presented in Tables 2 and 3, respectively. The "Label" column in these tables corresponds to the naming convention adopted by VASP. The checksum of each file listed in the tables is included in the accompanying supplement for verification purposes.

Each potential provided with the VASP package has two recommended plane-wave kinetic energy cut-off (E_{cut}) values, the smaller of which ensures the reliability of a calculation to within a well-defined error. Additionally, materials with more than one element type will have two or more sets of recommended E_{cut} values. In the AFLOW standard, the applied E_{cut} value is the largest found among the recommendations for all species involved in the calculation, increased by a factor of 1.4.

Table 2
Projector-Augmented Wavefunction (PAW) potentials, parameterized for the LDA, PW91, and PBE functionals, included in the AFLOW standard. The PAW-PBE combination is used as the default for ICSD, Binary Alloy and Heusler databases.

Element	Label	Element	Label	Element	Label
H	H	Se	Se	Gd[b]	Gd_3
He	He	Br	Br	Tb	Tb_3
Li	Li_sv	Kr	Kr	Dy	Dy_3
Be	Be_sv	Rb	Rb_sv	Ho	Ho_3
B	B_h	Sr	Sr_sv	Er	Er_3
C	C	Y	Y_sv	Tm	Tm
N	N	Zr	Zr_sv	Yb	Yb
O	O	Nb	Nb_sv	Lu	Lu
F	F	Mo	Mo_pv	Hf	Hf
Ne	Ne	Tc	Tc_pv	Ta	Ta_pv
Na	Na_pv	Ru	Ru_pv	W	W_pv
Mg	Mg_pv	Rh	Rh_pv	Re	Re_pv
Al	Al	Pd	Pd_pv	Os	Os_pv
Si	Si	Ag	Ag	Ir	Ir
P	P	Cd	Cd	Pt	Pt
S	S	In	In_d	Au	Au
Cl	Cl	Sn	Sn	Hg	Hg
Ar	Ar	Sb	Sb	Tl	Tl_d
K	K_sv	Te	Te	Pb	Pb_d
Ca	Ca_sv	I	I	Bi	Bi_d
Sc	Sc_sv	Xe	Xe	Po	Po
Ti	Ti_sv	Cs	Cs_sv	At	At
V	V_sv	Ba	Ba_sv	Rn	Rn
Cr	Cr_pv	La	La	Fr	Fr
Mn	Mn_pv	Ce	Ce	Ra	Ra
Fe	Fe_pv	Pr	Pr	Ac	Ac
Co	Co	Nd	Nd	Th	Th_s
Ni	Ni_pv	Pm	Pm	Pa	Pa
Cu	Cu_pv	Sm[a]	Sm	U	U
Zn	Zn	Sm[b]	Sm_3	Np	Np_s
Ga	Ga_h	Eu	Eu	Pu	Pu_s
As	As	Gd[a]	Gd		

[a] PBE potentials only.
[b] LDA and PW91 potentials only.

It is possible to evaluate the non-local parts of the potentials in real space, rather than in the more computationally intensive reciprocal space. This approach is prone to aliasing errors, and requires the optimization of real-space projectors if these are to be avoided. The real-space projection scheme is most appropriate for larger systems, e.g. surfaces, and is therefore not used in the construction of the databases found in the AFLOWLIB repository.

3.3. Fourier transform meshes

As mentioned previously, it is not practical to describe the precise default settings that are applied by the AFLOW standard in the specification of the Fourier transform meshes. We shall just note that they are defined in terms of the grid spacing along each of the reciprocal lattice vectors, $\vec{b}_i$. These are obtained from the set of real space lattice vectors, $\vec{a}_i$, via $[\vec{b}_1\vec{b}_2\vec{b}_3]^T = 2\pi[\vec{a}_1\vec{a}_2\vec{a}_3]^{-1}$. A distance in reciprocal space is then defined by $d_i = \|\vec{b}_i\|/n_i$, where the set of n_i are the number of grid points along each reciprocal lattice vector, and where the total number of points in the simulation is $n_1 \times n_2 \times n_3$.

The VASP package relies primarily on the so-called *dual grid technique*, which consists of two overlapping meshes with different coarseness. The least dense of the two is directly dependent on the applied plane-wave basis, E_{cut}, while the second is a finer mesh onto which the charge density is mapped. The AFLOW standard relies on placing sufficient points in the finer mesh such that wrap-around ("aliasing") errors are avoided. In terms of the quantity d_i, defined above, the finer grid is characterized by $d_i \approx 0.10$ Å^{-1}, while the coarse grid results in $d_i \approx 0.15$ Å^{-1}. These

Table 3
Ultra-Soft Pseudopotentials (USPP), parameterized for the LDA and PW91 functionals, included in the AFLOW standard.

Element	Label	Element	Label	Element	Label
H	H_soft	As	As	Tb	Tb_3
He	He	Se	Se	Dy	Dy_3
Li	Li_pv	Br	Br	Ho	Ho_3
Be	Be	Kr	Kr	Er	Er_3
B	B	Rb	Rb_pv	Tm	Tm
C	C	Sr	Sr_pv	Yb	Yb
N	N	Y	Y_pv	Lu	Lu
O	O	Zr	Zr_pv	Hf	Hf
F	F	Nb	Nb_pv	Ta	Ta
Ne	Ne	Mo	Mo_pv	W	W
Na	Na_pv	Tc	Tc	Re	Re
Mg	Mg_pv	Ru	Ru	Os	Os
Al	Al	Rh	Rh	Ir	Ir
Si	Si	Pd	Pd	Pt	Pt
P	P	Ag	Ag	Au	Au
S	S	Cd	Cd	Hg	Hg
Cl	Cl	In	In_d	Tl	Tl_d
Ar	Ar	Sn	Sn	Pb	Pb
K	K_pv	Sb	Sb	Bi	Bi
Ca	Ca_pv	Te	Te	Po	Po
Sc	Sc_pv	I	I	At	At
Ti	Ti_pv	Xe	Xe	Rn	Rn
V	V_pv	Cs	Cs_pv	Fr	Fr
Cr	Cr	Ba	Ba_pv	Ra	Ra
Mn	Mn	La	La	Ac	Ac
Fe	Fe	Ce	Ce	Th	Th_s
Co	Co	Pr	Pr	Pa	Pa
Ni	Ni	Nd	Nd	U	U
Cu	Cu	Pm	Pm	Np	Np_s
Zn	Zn	Sm	Sm_3	Pu	Pu_s
Ga	Ga_d	Eu	Eu		
Ge	Ge	Gd	Gd		

two values are approximate, as there is significant dispersion in these quantities across the various databases.

3.4. DFT+U corrections

Extended systems containing *d* and *f* block elements are often poorly represented within DFT due to the well known self interaction error (SIE) [27]. The influence that the SIE has on the energy gap of insulators has long been recognized, and several methods that account for it are available. These include the *GW* approximation [32], the rotationally invariant approach introduced by Dudarev [33] and Liechtenstein [34] (denoted here as DFT+U), as well as the recently developed ACBN0 pseudo-hybrid density functional [35].

The DFT+U approach is currently the best suited for high-throughput investigations, and is therefore included in the AFLOW standard for the entire ICSD database, and is also used for certain entries in the Heusler database containing the elements O, S, Se, and F. It is not used for the Binary Alloy database. This method has a significant dependence on parameters, as each atom is associated with two numbers, the screened Coulomb parameter, *U*, and the Stoner exchange parameter, *J*. These are usually reported as a single factor, combined via $U_{\text{eff}} = U - J$. The set of U_{eff} values associated with the *d* block elements [18,36] are presented in Table 4, to which the elements In and Sn have been added.

A subset of the *f*-block elements can be found among the systems included in the AFLOWLIB consortium databases. We are not aware of the existence of a systematic search for the best set of *U* and *J* parameters for the elements Nd, Sm, and Eu, so we have relied on an in-house parameterization [18] for those entries in the databases that contain them. The values used are reproduced in Table 5. Note that by construction the SIE correction must be applied to a pre-selected value of the ℓ-quantum number, and all elements listed in Table 4 correspond to $\ell = 2$, while those found in Table 5 correspond to $\ell = 3$.

Table 4
U_{eff} parameters applied to *d* orbitals.

Element	U_{eff}	Refs.	Element	U_{eff}	Refs.
Sc	2.9	[38]	W	2.2	[39]
Ti	4.4	[40]	Tc	2.7	[39]
V	2.7	[41]	Ru	3.0	[39]
Cr	3.5	[42]	Rh	3.3	[39]
Mn	4.0	[42]	Pd	3.6	[39]
Fe	4.6	[43]	Ag	5.8	[44]
Co	5.0	[41]	Cd	2.1	[45]
Ni	5.1	[41]	In	1.9	[45]
Cu	4.0	[42]	Sn	3.5	[46]
Zn	7.5	[45]	Ta	2.0	[39]
Ga	3.9	[47]	Re	2.4	[39]
Sn	3.5	[46]	Os	2.6	[39]
Nb	2.1	[39]	Ir	2.8	[39]
Mo	2.4	[39]	Pt	3.0	[39]
Ta	2.0	[46]	Au	4.0	

Table 5
U and *J* parameters applied to selected *f*-block elements.

Element	*U*	*J*	Refs.	Element	*U*	*J*	Refs.
La	8.1	0.6	[48]	Dy	5.6	0.0	[49]
Ce	7.0	0.7	[50]	Tm	7.0	1.0	[51]
Pr	6.5	1.0	[52]	Yb	7.0	0.67	[53]
Nd	7.2	1.0	[18]	Lu	4.8	0.95	[48]
Sm	7.4	1.0	[18]	Th	5.0	0.0	[54]
Eu	6.4	1.0	[18]	U	4.0	0.0	[55]
Gd	6.7	0.1	[56]				

The *U* and *J* values listed have only been applied to neutral systems, given that there are no entries in the AFLOWLIB repositories that contain determinations of charged states. The in-house parameterization [18] applicable to Nd, Sm, and Eu was performed by fitting their 4*f* levels to the corresponding experimental density of states obtained from X-ray Photoelectron Spectroscopy–Brems strahlung Isochromat Spectroscopy (XPS–BIS) data [37]. All other DFT+U parameters have been taken from the literature, and the corresponding citations are also contained in Tables 4 and 5.

3.5. Spin polarization

The first of the two `RELAX` calculations is always performed in a collinear spin-polarized fashion. The initial magnetic moments in this step are set to the number of atoms in the system, e.g. 1.0 μB/atom. If the magnetization resulting from the `RELAX1` step is found to be below 0.025 μB/atom, AFLOW economizes computational resources by turning spin polarization off in all ensuing calculations. Spin–orbit coupling is not used in the current AFLOW standard, since it is still too expensive to include in a HT framework.

3.6. Calculation methods and convergence criteria

Two nested loops are involved in the DFT calculations used by AFLOW in the construction of the databases. The inner loop contains routines that iteratively optimize the electronic degrees of freedom (EDOF), and features a number of algorithms that are concerned with diagonalizing the Kohn–Sham (KS) Hamiltonian at each iteration. The outer loop performs adjustments to the system geometry (ionic degrees of freedom, IDOF) until the forces acting on the system are minimized.

The convergence condition for each loop has been defined in terms of an energy difference, δE. If successive energies resulting from the completion of a loop are denoted as E_{i-1} and E_i, then convergence is met when the condition $\delta E \geqslant E_i - E_{i-1}$ is fulfilled. Note that E_i can either be the electronic energy resulting from the inner loop, or the configurational energy resulting from the outer loop. The electronic convergence criteria will be denoted as δE_{elec}, and the ionic criteria as δE_{ion}. The AFLOW standard relies on $\delta E_{elec} = 10^{-5}$ eV and $\delta E_{ion} = 10^{-4}$ eV for entries in the Elements database. All other databases include calculations performed with $\delta E_{elec} = 10^{-3}$ eV and $\delta E_{ion} = 10^{-2}$ eV.

Optimizations of the EDOF depend on sets of parameters that fall under three general themes: initial guesses, diagonalization methods, and charge mixing. The outer loop (optimizations of the IDOF) is concerned with the lattice vectors and the ionic positions, and is not as dependent on user input as the inner loops. These are described in the following paragraphs.

3.6.1. Electronic degrees of freedom

The first step in the process of optimizing the EDOF consists of choosing a trial charge density and a trial wavefunction. In the case of the non-BANDS-type calculations, the trial wavefunctions are initialized using random numbers, while the trial charge density is obtained from the superposition of atomic charge densities. The BANDS calculations are not self-consistent, and thus do not feature a charge density optimization. In these cases the charge density obtained from the previously performed STATIC calculation is used in the generation of the starting wavefunctions.

Two iterative methods are used for diagonalizing the KS Hamiltonian: the Davidson blocked scheme (DBS) [57,58], and the preconditioned residual minimization method–direct inversion in the iterative subspace (RMM–DIIS) [10]. Of the two, DBS is known to be the slower and more stable option. Additionally, the subspace rotation matrix is always optimized. These methods are applied in a manner that is dependent on the calculation type:

i. RELAX calculations. Geometry optimizations contain at least one determination of the system forces. The initial determination consists of 5 initial DBS steps, followed by as many RMM–DIIS steps as needed to fulfill the δE_{elec} condition. Later determinations of system forces are performed by a similar sequence, but only a single DBS step is applied at the outset of the process. Across all databases the minimum of number of electronic iterations for RELAX calculations is 2. The maximum number is set to 120 for entries in the ICSD, and 60 for all others.
ii. non-RELAX calculations. In STATIC or BANDS calculations, the diagonalizations are always performed using RMM–DIIS. The minimum number of electronic iterations performed during non-RELAX calculations is 2, and the maximum is 120.

If the number of iterations in the inner loop somehow exceed the limits listed above, the calculation breaks out of this loop, and the system forces and energy are determined. If the δE_{ion} convergence condition is not met the calculation re-enters the inner loop, and proceeds normally.

Charge mixing is performed via Pulay's method [59]. The implementation of this charge mixing approach in the VASP package depends on a series of parameters, of which all but the maximum ℓ-quantum number handled by the mixer have been left in their default state. This parameter is modified only in systems included in the ICSD database which contain the elements listed in Tables 4 and 5. In practical terms, the value applied in these cases is the maximum ℓ-quantum number found in the PAW potential, multiplied by 2.

3.6.2. Ionic degrees of freedom and lattice vectors

The RELAX calculation type contains determinations of the forces acting on the ions, as well as the full system stress tensor. The applied algorithm is the conjugate gradients (CG) approach [60], which depends on these quantities for the full optimization of the system geometry, i.e. the ionic positions, the lattice vectors, as well as modifications of the cell volume. The implementation of CG in VASP requires minimal user input, where the only independent parameter is the initial scaling factor which is always left at its default value. Convergence of the IDOF, as stated above, depends on the value for the δE_{ion} parameter, as applied across the various databases. The adopted E_{cut} (see discussion on "Potentials and basis set", Section 3.2) makes corrections for Pulay stresses unnecessary.

Forces acting on the ions and stress tensor are subjected to Harris–Foulkes [61] corrections. Molecular dynamics based relaxations are not performed in the construction of the databases found in the AFLOWLIB repository, so any related settings are not applicable to this work.

3.7. Output options

The reproduction of the results presented on the AFLOWLIB website also depends on a select few parameters that govern the output of the DFT package. The density of states plots are generated from the STATIC calculation. States are plotted with a range of −30 eV to 45 eV, and with a resolution of 5000 points. The band structures are plotted according to the paths of k-points generated for a BANDS calculation [18]. All bands found between −10 eV and 10 eV are included in the plots.

4. Conclusion

The AFLOW standard described here has been applied in the automated creation of the AFLOWLIB database of material properties in a consistent and reproducible manner. The use of standardized parameter sets facilitates the direct comparison of properties between different materials, so that specific trends can be identified to assist in the formulation of design rules for accelerated materials development. Following this AFLOW standard should allow materials science researchers to reproduce the results reported by the AFLOWLIB consortium, as well as to extend on the database and make meaningful comparisons with their own results.

Acknowledgments

We thank Dr. Kesong Yang for various technical discussions. We would like to acknowledge support by the DOD-ONR (N00014-13-1-0635, N00014-11-1-0136, N00014-09-1-0921), and CT, JJP and SC acknowledge support from the DOE (DE-AC02-05CH11231), specifically the Basic Energy Sciences program under Grant # EDCBEE. The AFLOWLIB consortium would like to acknowledge the Duke University Center for Materials Genomics and the CRAY corporation for computational support. C.O. acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGF1106401.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.commatsci.2015.07.019.

References

[1] S. Curtarolo, W. Setyawan, G.L.W. Hart, M. Jahnatek, R.V. Chepulskii, R.H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. Mehl, H.T. Stokes, D.O. Demchenko, D. Morgan, Comput. Mater. Sci. 58 (2012) 218–226, http://dx.doi.org/10.1016/j.commatsci.2012.02.005.

[2] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R.H. Taylor, L.J. Nelson, G.L.W. Hart, S. Sanvito, M. Buongiorno Nardelli, N. Mingo, O. Levy, Comput. Mater. Sci. 58 (2012) 227–235, http://dx.doi.org/10.1016/j.commatsci.2012.02.002.

[3] R.H. Taylor, F. Rose, C. Toher, O. Levy, K. Yang, M. Buongiorno Nardelli, S. Curtarolo, Comput. Mater. Sci. 93 (2014) 178–192, http://dx.doi.org/10.1016/j.commatsci.2014.05.014.

[4] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K.A. Persson, APL Mater. 1 (1) (2013) 011002, http://dx.doi.org/10.1063/1.4812323.

[5] S.P. Ong, W.D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V.L. Chevrier, K.A. Persson, G. Ceder, Comput. Mater. Sci. 68 (2013) 314–319, http://dx.doi.org/10.1016/j.commatsci.2012.10.028.

[6] Aflowlib.org, Ab-initio Electronic Structure Library Database, <http://www.aflowlib.org>.

[7] G. Kresse, J. Hafner, Phys. Rev. B 47 (1993) 558–561, http://dx.doi.org/10.1103/PhysRevB.47.558.

[8] G. Kresse, J. Hafner, Phys. Rev. B 49 (1994) 14251–14269, http://dx.doi.org/10.1103/PhysRevB.49.14251.

[9] G. Kresse, J. Furthmüller, Comput. Mater. Sci. 6 (1996) 15–50, http://dx.doi.org/10.1016/0927-0256(96)00008-0.

[10] G. Kresse, J. Furthmüller, Phys. Rev. B 54 (1996) 11169–11186, http://dx.doi.org/10.1103/PhysRevB.54.11169.

[11] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, J. Phys.: Condens. Matter 21 (39) (2009) 395502. <http://stacks.iop.org/0953-8984/21/i=39/a=395502>.

[12] C. Toher, J.J. Plata, O. Levy, M. de Jong, M. Asta, M. Buongiorno Nardelli, S. Curtarolo, Phys. Rev. B 90 (2014) 174107, http://dx.doi.org/10.1103/PhysRevB.90.174107.

[13] G. Bergerhoff, R. Hundt, R. Sievers, I.D. Brown, J. Chem. Inf. Comput. Sci. 23 (2) (1983) 66–69, http://dx.doi.org/10.1021/ci00038a003.

[14] A. Belsky, M. Hellenbrandt, V.L. Karen, P. Luksch, Acta Cryst. B58 (2002) 364–369.

[15] T.B. Massalski, H. Okamoto, P.R. Subramanian, L. Kacprzak (Eds.), Binary Alloy Phase Diagrams, American Society for Metals, Materials Park, OH, 1990.

[16] S. Curtarolo, D. Morgan, G. Ceder, Calphad 29 (2005) 163–211.

[17] P. Villars, K. Cenzual, J.L.C. Daams, F. Hulliger, T.B. Massalski, H. Okamoto, K. Osaki, A. Prince, S. Iwata, Crystal Impact, Pauling File. Inorganic Materials Database and Design System, Binaries edition., ASM International, Metal Park, OH, 2003.

[18] W. Setyawan, S. Curtarolo, Comput. Mater. Sci. 49 (2010) 299–312, http://dx.doi.org/10.1016/j.commatsci.2010.05.010.

[19] H.J. Monkhorst, J.D. Pack, Phys. Rev. B 13 (12) (1976) 5188–5192, http://dx.doi.org/10.1103/PhysRevB.13.5188.

[20] P.E. Blöchl, O. Jepsen, O.K. Andersen, Phys. Rev. B 49 (23) (1994) 16223–16233, http://dx.doi.org/10.1103/PhysRevB.49.16223.

[21] M. Methfessel, A.T. Paxton, Phys. Rev. B 40 (1989) 3616–3621, http://dx.doi.org/10.1103/PhysRevB.40.3616.

[22] D. Vanderbilt, Phys. Rev. B 41 (1990) 7892.

[23] G. Kresse, J. Hafner, J. Phys.: Condens. Matter 6 (1994) 8245–8257.

[24] P.E. Blöchl, Phys. Rev. B 50 (1994) 17953–17979.

[25] G. Kresse, D. Joubert, Phys. Rev. B 59 (1999) 1758.

[26] D.M. Ceperley, B.J. Alder, Phys. Rev. Lett. 45 (1980) 566–569, http://dx.doi.org/10.1103/PhysRevLett.45.566.

[27] J.P. Perdew, A. Zunger, Phys. Rev. B 23 (1981) 5048–5079, http://dx.doi.org/10.1103/PhysRevB.23.5048.

[28] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, C. Fiolhais, Phys. Rev. B 46 (1992) 6671–6687, http://dx.doi.org/10.1103/PhysRevB.46.6671.

[29] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, C. Fiolhais, Phys. Rev. B 48 (1993), http://dx.doi.org/10.1103/PhysRevB.48.4978.2. pp. 4978–4978.

[30] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865–3868.

[31] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 78 (1997) 1396.

[32] L. Hedin, Phys. Rev. 139 (1965) A796–A823, http://dx.doi.org/10.1103/PhysRev.139.A796.

[33] S.L. Dudarev, G.A. Botton, S.Y. Savrasov, C.J. Humphreys, A.P. Sutton, Phys. Rev. B 57 (1998) 1505–1509, http://dx.doi.org/10.1103/PhysRevB.57.1505.

[34] A.I. Liechtenstein, V.I. Anisimov, J. Zaanen, Phys. Rev. B 52 (1995) R5467–R5470, http://dx.doi.org/10.1103/PhysRevB.52.R5467.

[35] L.A. Agapito, S. Curtarolo, M. Buongiorno Nardelli, Phys. Rev. X 5 (2015) 011006, http://dx.doi.org/10.1103/PhysRevX.5.011006.

[36] S. Wang, Z. Wang, W. Setyawan, N. Mingo, S. Curtarolo, Phys. Rev. X 1 (2011) 021012, http://dx.doi.org/10.1103/PhysRevX.1.021012.

[37] J.K. Lang, Y. Baer, P.A. Cox, J. Phys. F: Met. Phys. 11 (1981) 121.

[38] H. Jiang, R.I. Gomez-Abal, P. Rinke, M. Scheffler, Phys. Rev. B 82 (2010) 045108, http://dx.doi.org/10.1103/PhysRevB.82.045108.

[39] I.V. Solovyev, P.H. Dederichs, V.I. Anisimov, Phys. Rev. B 50 (1994) 16861.

[40] D.D. Cuong, B. Lee, K.M. Choi, H.-S. Ahn, S. Han, J. Lee, Phys. Rev. Lett. 98 (2007) 115503.

[41] W.E. Pickett, S.C. Erwin, E.C. Ethridge, Phys. Rev. B 58 (1998) 1201.

[42] L. Wang, T. Maxisch, G. Ceder, Phys. Rev. B 73 (2006) 195107.

[43] M. Cococcioni, S. de Gironcoli, Phys. Rev. B 71 (2005) 035105.

[44] N. Umezawa, O. Shuxin, J. Ye, Phys. Rev. B 83 (2011) 035202, http://dx.doi.org/10.1103/PhysRevB.83.035202.

[45] P. Erhart, K. Albe, A. Klein, Phys. Rev. B 73 (2006) 205203.

[46] A.K. Singh, A. Janotti, M. Scheffler, C.G. Van de Walle, Phys. Rev. Lett. 101 (2008) 055502, http://dx.doi.org/10.1103/PhysRevLett.101.055502. sn d-orbital U = 3.5 eV.

[47] A. Janotti, D. Segev, C.G.V. de Walle, Phys. Rev. B 74 (2006) 045202.

[48] D. Wegner, A. Bauer, Y.M. Koroteev, G. Bihlmayer, E.V. Chulkov, P.M. Echenique, G. Kaindl, Phys. Rev. B 73 (2006) 115403.

[49] M. Luszczek, Physica C 469 (21) (2009) 1892–1897, http://dx.doi.org/10.1016/j.physc.2009.06.003.

[50] Y. Jiang, J.B. Adams, M. van Schilfgaarde, J. Chem. Phys. 123 (2005) 064701.

[51] V.N. Antonov, B.N. Harmon, A.N. Yaresko, Phys. Rev. B 63 (2001) 205112.

[52] H. Harima, J. Magn. Magn. Mater. 226 (2010) 83.

[53] T. Jeong, J. Phys.: Condens. Matter 18 (2006) 6769.

[54] C. Sevik, T. Çağın, Phys. Rev. B 80 (2009) 014108, http://dx.doi.org/10.1103/PhysRevB.80.014108.

[55] M. Sanati, R.C. Albers, T. Lookman, A. Saxena, Phys. Rev. B 84 (2011) 014116, http://dx.doi.org/10.1103/PhysRevB.84.014116.

[56] B.N. Harmon, V.P. Antropov, A.I. Liechtenstein, I.V. Solovyev, V.I. Anisimov, J. Phys. Chem. Solids 56 (1995) 1521.

[57] B. Liu, in: C. Moler, I. Shavitt (Eds.), Numerical Algorithms in Chemistry: Algebraic Methods, Lawrence Berkeley Laboratory, University of California, 1978, pp. 49–53.

[58] E.R. Davidson, in: G.H.F. Diercksen, S. Wilson (Eds.), NATO Science Series C: Methods in Computational Molecular Physics, vol. 3, D. Reidel Publishing Company, 1983, pp. 95–113.

[59] P. Pulay, Chem. Phys. Lett. 73 (1980) 393–398, http://dx.doi.org/10.1016/0009-2614(80)80396-4.

[60] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, 1992, ISBN 0521431085.

[61] J. Harris, Phys. Rev. B 31 (1985) 1770–1779, http://dx.doi.org/10.1103/PhysRevB.31.1770.