User's Guide

to

OCCQP

Version 0.39 for MATLAB® v5.2

Jeffrey A. Foley

Department of Geology

Miami University

Oxford, OH 45056 USA

Contents

Introduction 3

Basic Use 3

Input files 3

Example 4

Running OCCQP 5

Output residuals 7

Advanced Use 9

Testing other site assignments 11

Changing weights 12

The <File> Menu 14

Chemical Aggregation 15

Tips & Trouble Shooting 15

Appendix: Bond valence constants used in OccQP 16

References 18

System Requirements and Program Availability 18

Introduction

OCCQP, written in the MATLAB® programming language, uses quadratic programming (QP) to optimize site occupancies in crystals; Wright *et al*. (2000) provide details of the problem and the nomenclature used in this manual. All site assignments are based on x-ray structural and chemical data.

This manual has two parts: the first section explains the basic use of OCCQP, and the second section details how the user generates output files, bypasses stored constants, and tests the viability of any site assignment.

Basic Use

Input files. Each separate problem will have its own input file defining the following objects: having these variables defined:

```
Name  = 'Optional reference information';
Atoms ={ 'Atomic species' };
Cobs  =[ Chemical content in atoms per unit cell ];
Sites ={ 'Site names' };
Qobs  =[ Electron population of each site ];
C     =[ Rank of each site ];
Coord =[ Coordination number of each site ];
dA    ={ Bond length   'Anion' };
```

The *Name* variable is optional, but all others are essential. Input files must be flat text files (i.e., with no word-processor formatting) with names of the form "*filename*.m". Such files can be created with the MATLAB® text editor: start the MATLAB environment

(e.g., double-click on the MATLAB icon) and select the "File ⇒ New ⇒ M-file" pull-down menu item. Type in the input data (see the example below) and save the file with the desired filename.

Example 1. Dravite, Hawthorne *et al.* (1993). In this example we examine how four atomic species [Al, Fe(III), Fe(II), Mg] distribute between the *Y* and *Z* sites of a dravite tourmaline. First we create the input file "drav1.m":

```
Name  ='Dravite Tourmaline (Hawthorne et al., 1993, …
       AM 78,265)';

Atoms ={ 'Al'  'Fe(III)' 'Fe(II)' 'Mg' };
Cobs  =[6.075   0.560     0.051  2.211 ];

Sites ={  'Y'   'Z'   };
Qobs  =[  15.15 13.03 ];
C     =[   3     6    ];
Coord =[   6     6    ];

%  Anything to the right of '%' is a comment.
%      d    anion
dA={ 1.953   'O' % begin site Y
     2.002   'O'
     2.002   'O'
     2.117   'O'
     1.989   'O'
     1.989   'O'
     2.002   'O' % begin site Z
     1.915   'O'
     1.960   'O'
     1.908   'O'
     1.931   'O'
     1.900   'O' };
```

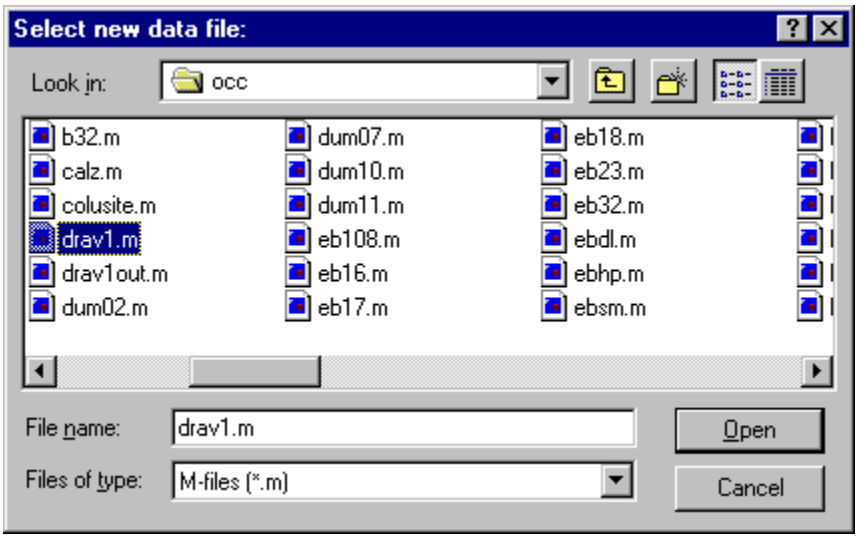
Atomic species have single quotes around them, and must be one of the species listed in Appendix 1. The observed chemistries, *Cobs*, must be in the same order as *Atoms*. Similar to *Atoms*, entries in *Sites* must have single quotes around them. The electron populations, *Qobs*, the site rank *C*, and coordination number, *Coord*, must be listed in the same order as *Sites*. The array of bond lengths and atoms bonded to the sites, *dA*, must also be listed in that order. Finally, each array definition should end with a semicolon to suppress extraneous screen output of the array.

Running OccQP. To run OccQP, start MATLAB® and make sure the file "occqp.m" is in the MATLAB search path. In the example below, this is accomplished by moving into the directory C:\Matlab\Research containing "occqp.m". Next call OccQP by entering its name at the MATLAB command line:

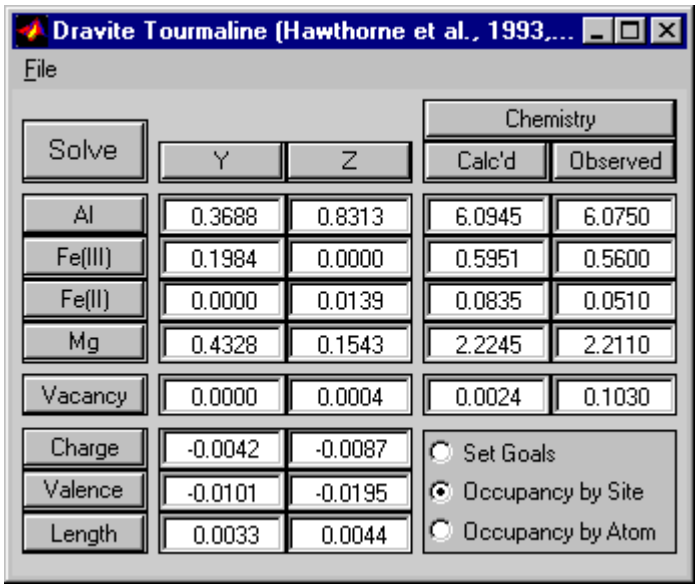
```
>> cd c:\matlab\research

>> occqp
```

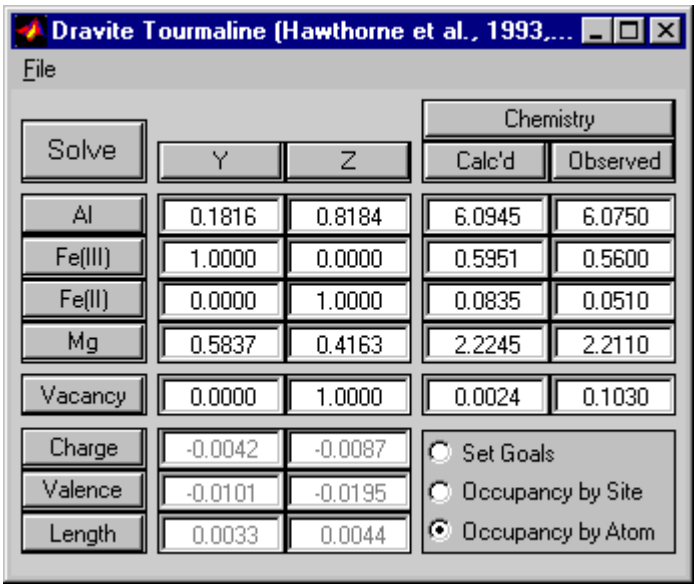
This command brings up a selection menu:



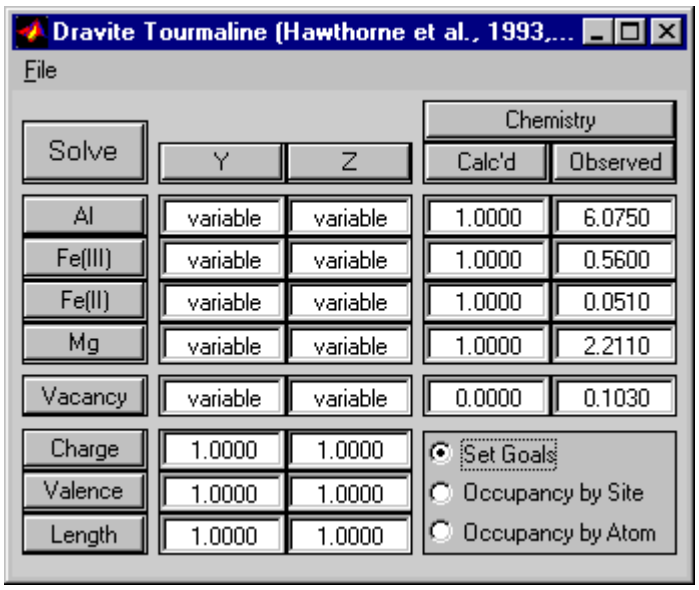
To select the previously created dravite input file (see p. 4), double-click on "drav1.m" to display the default site assignments:



The default optimum occupancies are calculated using unit weights on chemistry, charge, valence, and average bond length. Choose *Occupancy by Atom* to show how each atom is apportioned among the sites:



The weights and goals (click the *Set Goals* toggle) are the defaults (unit weight, all atomic species are allowed into either site, and vacancies are allowed):



Output residuals. The observed chemistry is simply the *Cobs* array specified in the input file. The observed vacancy is the difference between the sum of ranks and the sum of chemistries:

$$\text{observed vacancies} = \sum_{j}^{m} C^{(j)} - \sum_{k}^{n} C_k^{obs} .$$

See Wright et al. (2000) for nomenclature. This value could be negative, meaning that (according to the chemistry) there are more atoms than available sites. The values labeled *Charge*, *Valence*, and *Length* are the following residuals for each site j:

$$\sum_{k}^{n} x_k^{(j)} Q_k - Q_{obs}^{(j)},$$

$$\sum_{k=1}^{n} x_k^{(j)} Z_k - \sum_{k=1}^{n} x_k^{(j)} V_k^{(j)} ,$$

$$\sum_{k=1}^{n} x_k^{(j)} D_k^{(j)} - \sum_{k=1}^{n} x_k^{(j)} \hat{d}^{(j)} .$$

Advanced Use

For this section we use the same dravite, but consider five cation sites (X, Y, Z, T, and B) and all twelve atomic species. Make the following ".m" file, calling it “drav2.m”; note that to continue long lines of input simply use ellipses (…):

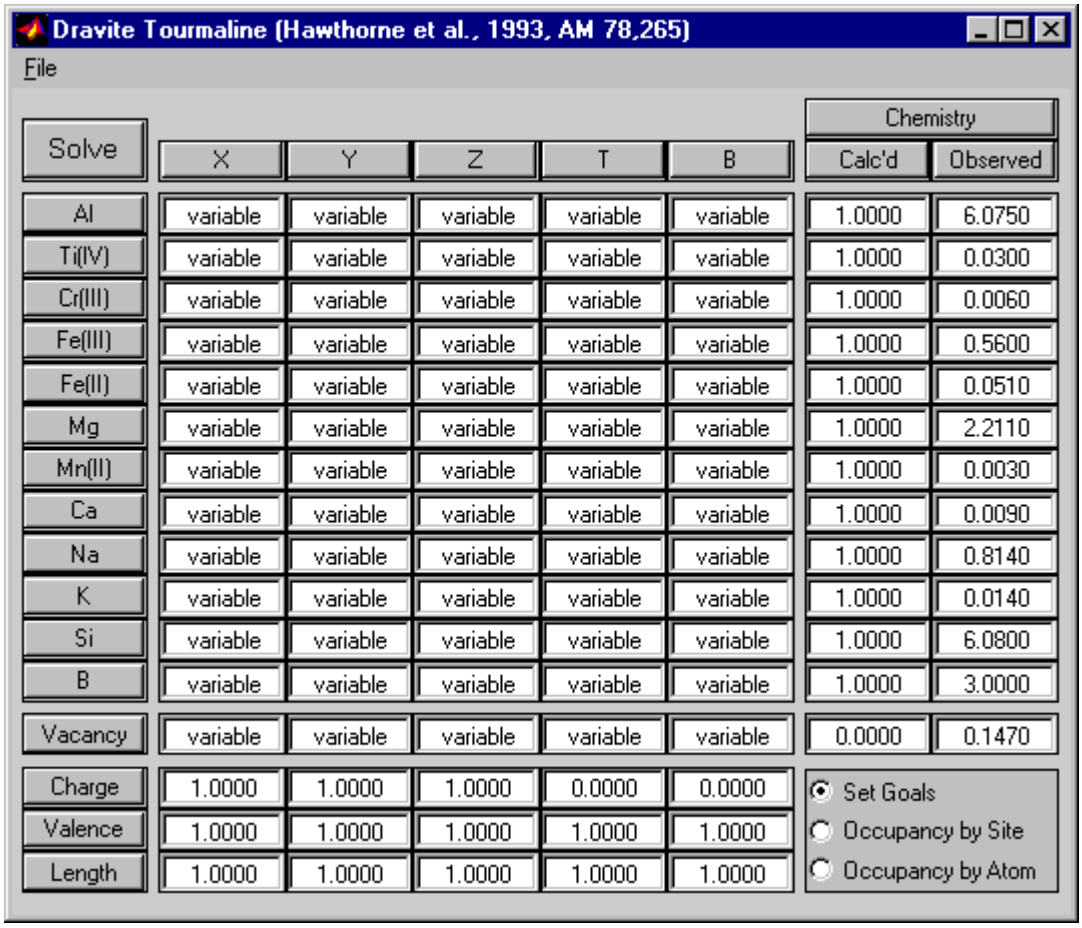
```
Name  ='Dravite Tourmaline (Hawthorne et al., 1993,…
       AM 78,265)';

Atoms = {'Al'     'Ti(IV)' 'Cr(III)' 'Fe(III)'…
         'Fe(II)' 'Mg'     'Mn(II)'  'Ca'…
         'Na'     'K'      'Si'      'B'};
Cobs  = [ 6.075    0.030    0.006     0.560…
          0.051    2.211    0.003     0.009…
          0.814    0.014    6.080     3];

Sites = {      'X'       'Y'       'Z'       'T'       'B' };
Qobs  = [    9.27     15.15     13.03         0         0  ];
C     = [       1         3         6         6         3  ];
Coord = [       9         6         6         4         3  ];

dA = {  2.5040  'O'    %'X'
        2.5040  'O'    %'X'
        2.5040  'O'    %'X'
        2.8170  'O'    %'X'
        2.8170  'O'    %'X'
        2.8170  'O'    %'X'
        2.7460  'O'    %'X'
        2.7460  'O'    %'X'
        2.7460  'O'    %'X'
        1.9530  'O'    %'Y'
        2.0020  'O'    %'Y'
        2.0020  'O'    %'Y'
        2.1170  'O'    %'Y'
        1.9890  'O'    %'Y'
        1.9890  'O'    %'Y'
        2.0020  'O'    %'Z'
        1.9150  'O'    %'Z'
        1.9600  'O'    %'Z'
        1.9080  'O'    %'Z'
        1.9310  'O'    %'Z'
        1.9000  'O'    %'Z'
        1.6260  'O'    %'T'
        1.6450  'O'    %'T'
        1.6060  'O'    %'T'
        1.6050  'O'    %'T'
        1.3726  'O'    %'B'
        1.3744  'O'    %'B'
        1.3744  'O' }; %'B'
```

Start OccQP and select “drav2.m”. The default weighting will give meaningless results because no values were reported for *Qobs* at the T and B sites. Therefore, select *Set Goals* and change the default weights on the *Charge* residuals to zero for these sites. This is accomplished by double-clicking on the desired entries and typing ‘0’.



The displayed *Charge* residual is actually an estimate of *Qobs*:

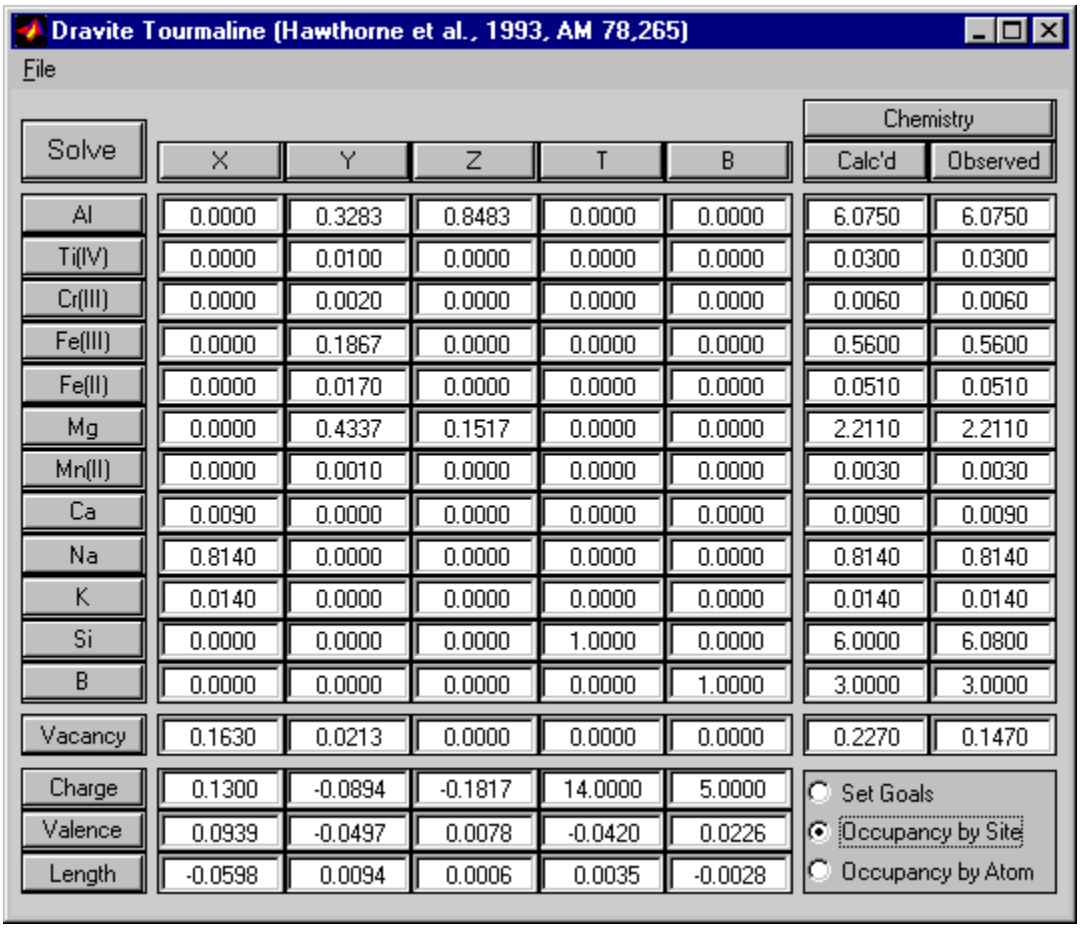
Dravite Tourmaline (Hawthorne et al., 1993, AM 78,265)

File

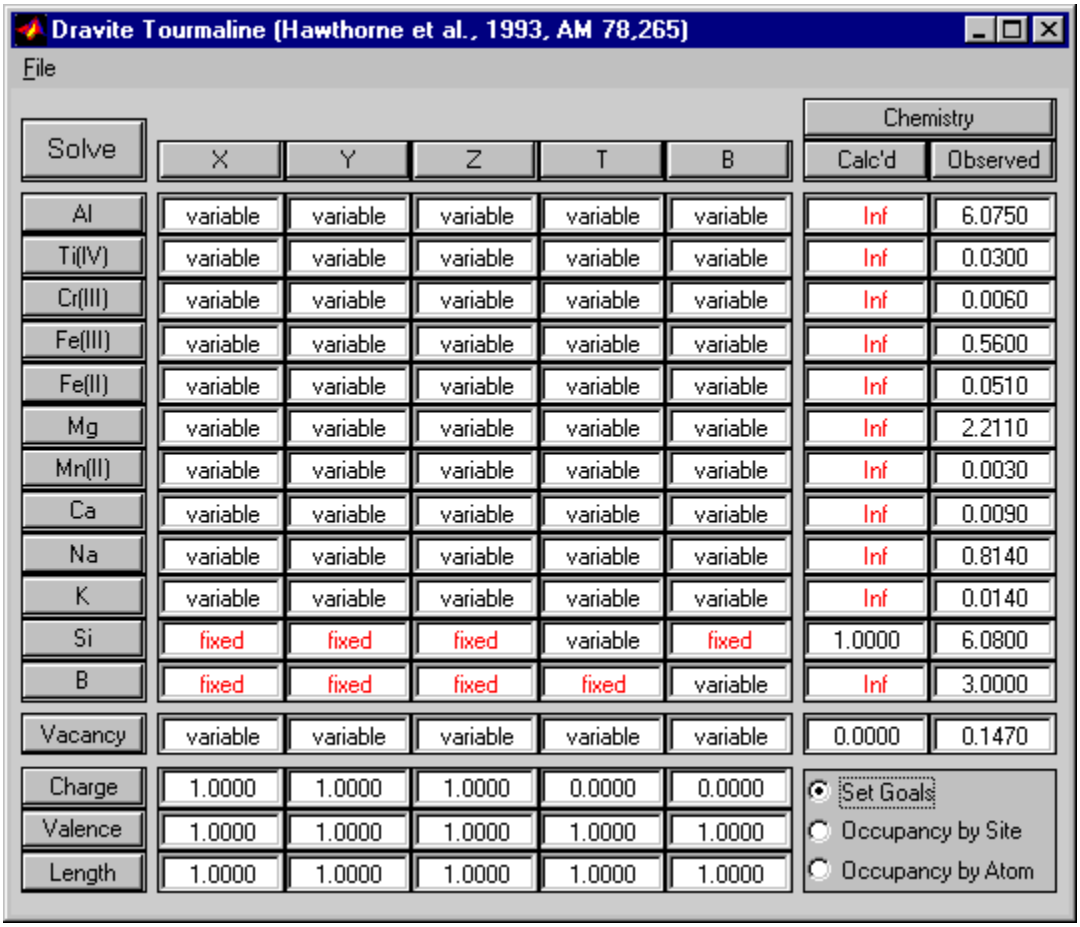
Solve	X	Y	Z	T	B	Chemistry: Calc'd	Chemistry: Observed
Al	0.0000	0.3657	0.8265	0.0000	0.0000	6.0559	6.0750
Ti(IV)	0.0000	0.0105	0.0000	0.0000	0.0000	0.0314	0.0300
Cr(III)	0.0000	0.0000	0.0020	0.0000	0.0000	0.0117	0.0060
Fe(III)	0.0000	0.1877	0.0011	0.0000	0.0000	0.5699	0.5600
Fe(II)	0.0000	0.0000	0.0099	0.0000	0.0000	0.0596	0.0510
Mg	0.0000	0.4191	0.1550	0.0000	0.0000	2.1874	2.2110
Mn(II)	0.0000	0.0028	0.0000	0.0000	0.0000	0.0084	0.0030
Ca	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0090
Na	0.7652	0.0065	0.0000	0.0000	0.0000	0.7848	0.8140
K	0.0450	0.0000	0.0000	0.0000	0.0000	0.0450	0.0140
Si	0.0000	0.0077	0.0055	0.9932	0.0166	6.0651	6.0800
B	0.0000	0.0000	0.0000	0.0068	0.9834	2.9911	3.0000
Vacancy	0.1898	-0.0000	0.0000	0.0000	0.0000	0.1898	0.1470
Charge	0.0018	-0.0070	-0.0141	13.9387	5.1494		
Valence	0.0470	-0.0046	-0.0092	-0.0352	-0.0093		
Length	-0.0467	0.0036	0.0031	0.0025	-0.0004		

- ○ Set Goals
- ◉ Occupancy by Site
- ○ Occupancy by Atom

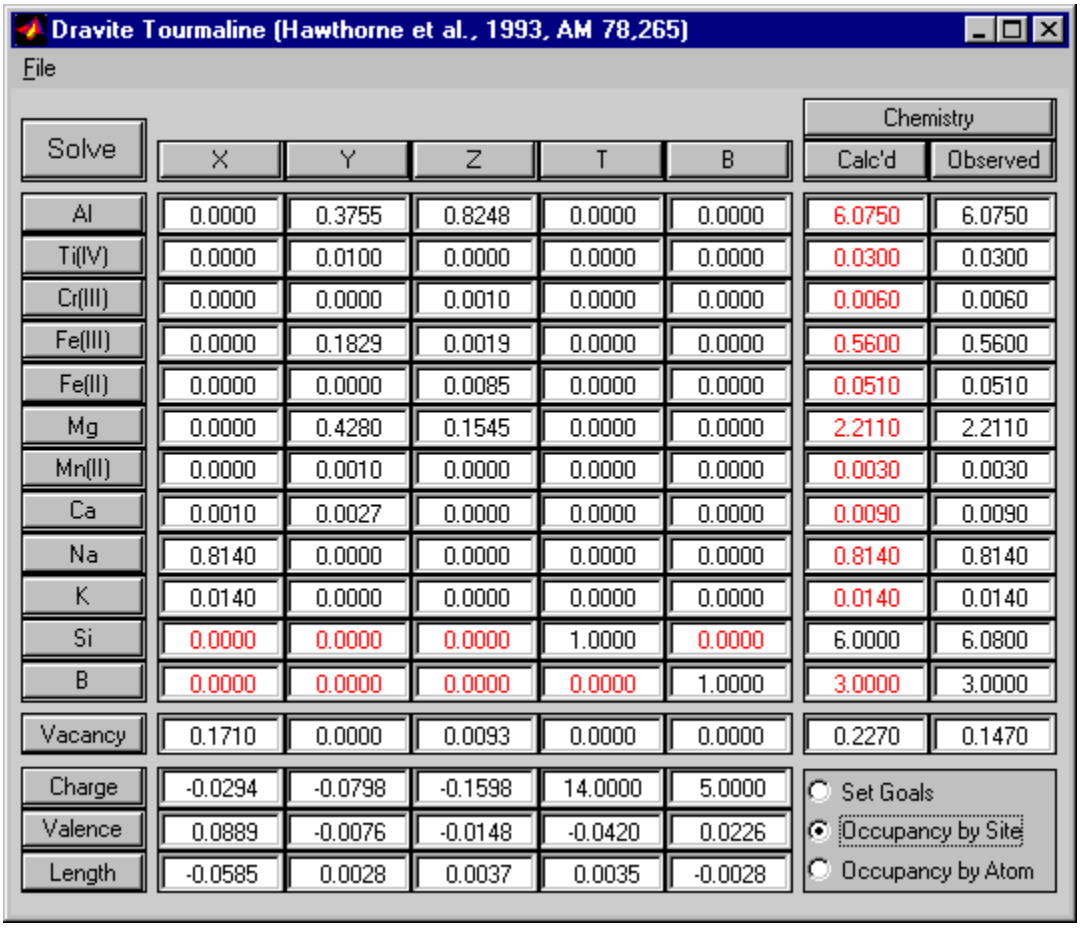
Testing other site assignments. OccQP can be used to determine residuals in published site-assignments. Simply double-click on any occupancy box in the output, and input the desired value. The residuals are recalculated using the input site occupancies. In the example below, the site assignments made by Hawthorne *et al.* (1993) are input and the residuals calculated:



Changing Weights. By changing weights, OccQP can perform site assignment calculations wherein the user makes judgements on the validity of each observation. In addition, site assignments can be input by the user. In the example below we fix all the Si in the T site, and all the B in the B site, as suggested by crystal chemical intuition. Infinite weights (inf) are placed on the chemistry equations in order to honor the analyzed chemistry exactly in the optimization.



The optimal solution corresponding to these settings is shown on the following page.



Dravite Tourmaline (Hawthorne et al., 1993, AM 78,265)

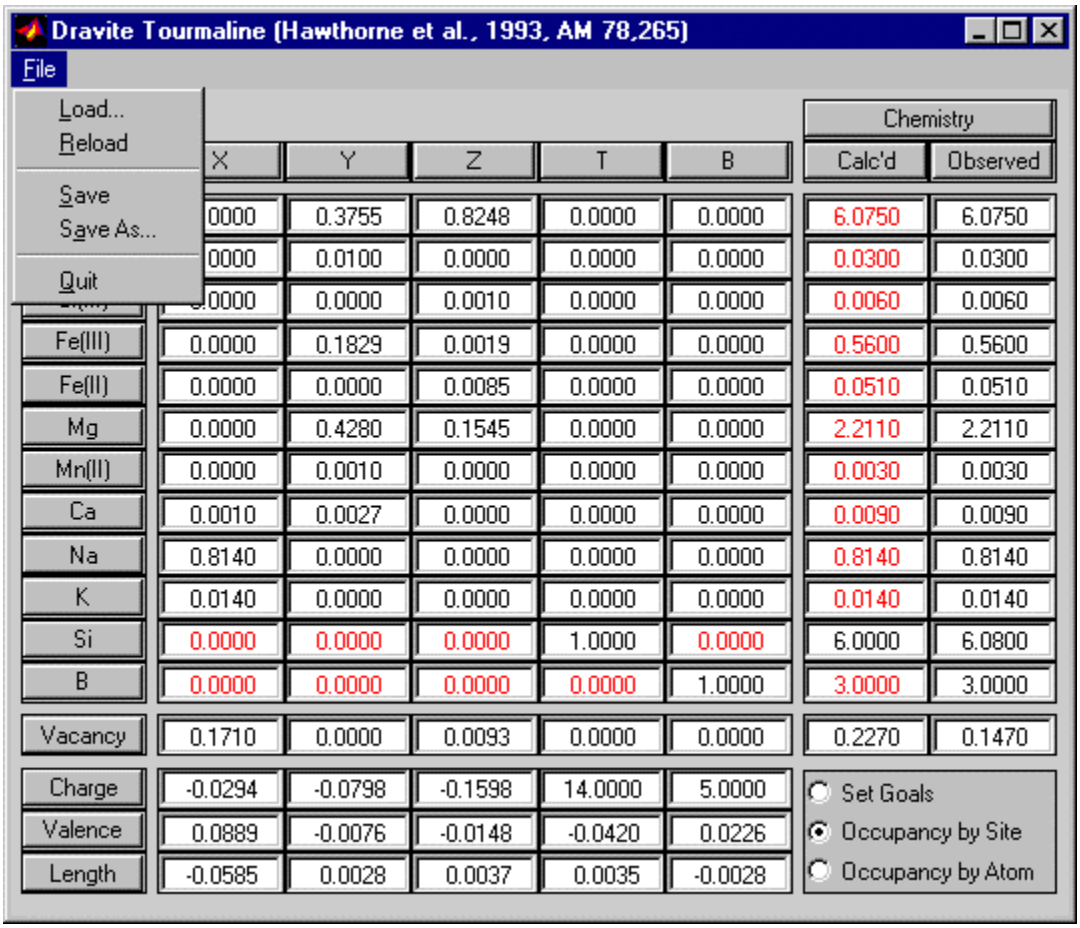
File

Solve	X	Y	Z	T	B	Chemistry Calc'd	Chemistry Observed
Al	0.0000	0.3755	0.8248	0.0000	0.0000	6.0750	6.0750
Ti(IV)	0.0000	0.0100	0.0000	0.0000	0.0000	0.0300	0.0300
Cr(III)	0.0000	0.0000	0.0010	0.0000	0.0000	0.0060	0.0060
Fe(III)	0.0000	0.1829	0.0019	0.0000	0.0000	0.5600	0.5600
Fe(II)	0.0000	0.0000	0.0085	0.0000	0.0000	0.0510	0.0510
Mg	0.0000	0.4280	0.1545	0.0000	0.0000	2.2110	2.2110
Mn(II)	0.0000	0.0010	0.0000	0.0000	0.0000	0.0030	0.0030
Ca	0.0010	0.0027	0.0000	0.0000	0.0000	0.0090	0.0090
Na	0.8140	0.0000	0.0000	0.0000	0.0000	0.8140	0.8140
K	0.0140	0.0000	0.0000	0.0000	0.0000	0.0140	0.0140
Si	0.0000	0.0000	0.0000	1.0000	0.0000	6.0000	6.0800
B	0.0000	0.0000	0.0000	0.0000	1.0000	3.0000	3.0000
Vacancy	0.1710	0.0000	0.0093	0.0000	0.0000	0.2270	0.1470
Charge	-0.0294	-0.0798	-0.1598	14.0000	5.0000		
Valence	0.0889	-0.0076	-0.0148	-0.0420	0.0226		
Length	-0.0585	0.0028	0.0037	0.0035	-0.0028		

○ Set Goals
◉ Occupancy by Site
○ Occupancy by Atom

(Optimal solution using weights shown in the preceding figure.)

The <File> menu. Load and Reload. *Load* works identically to the command "occqp" from the MATLAB® screen. Using *Reload* restarts the current file. This command is useful when the program "hangs up" or fails.



Save and Save As. These commands provide a quick way of recording the weighting scheme and constraints used for a calculation, as well as the results of an optimization. A saved file also contains the values used for *Q*, *Z*, and *R*, allowing the user to change these values.

Chemical aggregation. The optional variable *Caggr* allows grouping of two or more species into one "aggregated" species, such as aggregating Fe(II) and Fe(III) as total Fe. When the *Caggr* line is added to the input file, atoms with like values of *Caggr* will be constrained to the total value of *Cobs*, but the optimization process will apportion the atoms between the species to minimize residuals. For example, with addition of the *Caggr* line to the input below,

```
Atoms ={ 'Na'    'K'    'Ca'    'Al'    'Mg'    'Fe(II)'   'Fe(III)' };
Q     =  [ 11     19     20      13      12       26         26     ];
Z     =  [  1      1      2       3       2        2          3     ];
Cobs  =  [ 0.814  0.014  0.009   6.075   2.211    0.051      0.56   ];
Caggr =  [  1       2      3       4       5        6          6     ];
```

the amount of Fe(II) and Fe(III) will be varied, with the *Cobs* goal of 0.611 *total* Fe atoms. Such a utility is useful when total Fe is analyzed but no oxidation data are available: Fe will be apportioned between the two valences to best meet the goals.

Tips & Trouble Shooting

Making new input files. The most frequent problem in building new input files is naming the atoms incorrectly. Atomic symbols must be enclosed between single quotes, and atoms with more than one possible valence state must have the valence expressed as a Roman numeral enclosed in parentheses. Allowed atom nomenclature is given in the *Appendix 1*.

It is not necessary to end each line with a semicolon (;), but doing so reduces extraneous output to the screen.

Appendix 1

Values of Q, Z, and R as stored in QCCQP.m. R values from Brese and O'Keeffe (1991).

Cation	*Q*	*Z*	*R(O)*	*R(F)*	*R(Cl)*	*R(S)*
'Ac'	89	3	2.24	2.13	2.63	nan
'Ag'	47	1	1.805	1.80	2.09	2.15
'Al'	13	3	1.651	1.545	2.03	2.13
'Am'	95	3	2.11	2.00	2.48	nan
'As(III)'	33	3	1.789	1.70	2.16	2.26
'As(V)'	33	5	1.767	1.62	2.14	2.26
'Au'	79	3	1.833	1.81	2.17	2.03
'B'	5	3	1.371	1.31	1.74	1.82
'Ba'	56	2	2.29	2.19	2.69	2.77
'Be'	4	2	1.381	1.28	1.76	1.83
'Bi(III)'	83	3	2.09	1.99	2.48	2.55
'Bi(V)'	83	5	2.06	1.97	2.44	2.55
'Bk'	97	3	2.08	1.96	2.46	nan
'Br'	35	7	1.81	1.72	2.19	nan
'C'	6	4	1.39	1.32	1.76	1.82
'Ca'	20	2	1.967	1.842	2.37	2.45
'Cd'	48	2	1.904	1.811	2.23	2.29
'Ce(III)'	58	3	2.151	2.036	2.52	2.62
'Ce(IV)'	58	4	2.028	1.995	2.41	2.62
'Cf'	98	3	2.07	1.95	2.45	nan
'Cl'	17	7	1.632	1.55	2.00	nan
'Cm'	96	3	2.23	2.12	2.62	nan
'Co(II)'	27	2	1.692	1.64	2.01	2.06
'Co(III)'	27	3	1.70	1.62	2.05	2.06
'Cr(II)'	24	2	1.73	1.67	2.09	2.18
'Cr(III)'	24	3	1.724	1.64	2.08	2.18
'Cr(VI)'	24	6	1.794	1.74	2.12	2.18
'Cs'	55	1	2.42	2.33	2.79	2.89
'Cu(I)'	29	1	1.593	1.6	1.85	1.86
'Cu(II)'	29	2	1.679	1.60	2.00	1.86
'Dy'	66	3	2.036	1.922	2.41	2.47
'Er'	68	3	2.010	1.906	2.39	2.46
'Eu(II)'	63	2	2.147	2.04	2.53	2.53
'Eu(III)'	63	3	2.076	1.961	2.455	2.53
'Fe(II)'	26	2	1.734	1.65	2.06	2.16
'Fe(III)'	26	3	1.759	1.67	2.09	2.16
'Ga'	31	3	1.730	1.62	2.07	2.17
'Gd'	64	3	2.065	1.95	2.445	2.53
'Ge'	32	4	1.748	1.66	2.14	2.22
'H'	1	1	0.95	0.92	1.28	1.35
'Hf'	72	4	1.923	1.85	2.30	2.39
'Hg(I)'	80	1	1.90	1.81	2.28	2.32
'Hg(II)'	80	2	1.93	1.90	2.25	2.32
'Ho'	67	3	2.023	1.908	2.401	2.48
'I(V)'	53	5	2.00	1.90	2.38	nan
'I(VII)'	53	7	1.93	1.83	2.31	nan
'In'	49	3	1.902	1.79	2.28	2.36
'Ir'	77	5	1.916	1.82	2.30	2.38
'K'	19	1	2.13	1.99	2.52	2.59
'La'	57	3	2.172	2.057	2.545	2.64
'Li'	3	1	1.466	1.360	1.91	1.94

'Lu'	71	3	1.971	1.876	2.361	2.43
'Mg'	12	2	1.693	1.581	2.08	2.18
'Mn(II)'	25	2	1.790	1.698	2.13	2.20
'Mn(III)'	25	3	1.760	1.66	2.14	2.20
'Mn(IV)'	25	4	1.753	1.71	2.13	2.20
'Mn(VII)'	25	7	1.79	1.72	2.17	2.20
'Mo'	42	6	1.907	1.81	2.28	2.35
'N(III)'	7	3	1.361	1.37	1.75	nan
'N(V)'	7	5	1.432	1.36	1.80	nan
'Na'	11	1	1.80	1.677	2.15	2.28
'Nb'	41	5	1.911	1.87	2.27	2.37
'Nd'	60	3	2.117	2.008	2.492	2.59
'Ni'	28	2	1.654	1.599	2.02	2.04
'Os'	76	4	1.811	1.72	2.19	nan
'P'	15	5	1.604	1.521	1.99	2.11
'Pb(II)'	82	2	2.112	2.03	2.53	2.55
'Pb(IV)'	82	4	2.042	1.94	2.43	2.55
'Pd'	46	2	1.792	1.74	2.05	2.10
'Pr'	59	3	2.135	2.022	2.50	2.60
'Pt(II)'	78	2	1.768	1.68	2.05	2.08
'Pt(IV)'	78	4	1.879	1.759	2.17	2.08
'Pu'	94	3	2.11	2.00	2.48	nan
'Rb'	37	1	2.26	2.16	2.65	2.70
'Re'	75	7	1.97	1.86	2.23	2.37
'Rh'	45	3	1.791	1.71	2.17	2.15
'Ru'	44	4	1.834	1.74	2.21	2.16
'S(IV)'	16	4	1.644	1.60	2.02	2.07
'S(VI)'	16	6	1.624	1.56	2.03	2.07
'Sb(III)'	51	3	1.973	1.90	2.35	2.45
'Sb(V)'	51	5	1.942	1.80	2.30	2.45
'Sc'	21	3	1.849	1.76	2.23	2.32
'Se(IV)'	34	4	1.811	1.73	2.22	2.25
'Se(VI)'	34	6	1.788	1.69	2.16	2.25
'Si'	14	4	1.624	1.58	2.03	2.13
'Sm'	62	3	2.088	1.977	2.466	2.55
'Sn(II)'	50	2	1.984	1.925	2.36	2.45
'Sn(IV)'	50	4	1.905	1.84	2.28	2.45
'Sr'	38	2	2.118	2.019	2.51	2.59
'Ta'	73	5	1.920	1.88	2.30	2.39
'Tb'	65	3	2.049	1.936	2.427	2.51
'Te(IV)'	52	4	1.977	1.87	2.37	2.45
'Te(VI)'	52	6	1.917	1.82	2.30	2.45
'Th'	90	4	2.167	2.07	2.55	2.64
'Ti(III)'	22	3	1.791	1.723	2.17	2.24
'Ti(IV)'	22	4	1.815	1.76	2.19	2.24
'Tl(I)'	81	1	2.172	2.15	2.56	2.63
'Tl(III)'	81	3	2.003	1.88	2.32	2.63
'Tm'	69	3	2.000	1.842	2.38	2.45
'U(IV)'	92	4	2.112	2.034	2.48	2.56
'U(VI)'	92	6	2.075	1.966	2.46	2.56
'V(III)'	23	3	1.743	1.702	2.19	2.23
'V(IV)'	23	4	1.784	1.70	2.16	2.23
'V(V)'	23	5	1.803	1.71	2.16	2.23
'W'	74	6	1.921	1.83	2.27	2.39
'Y'	39	3	2.014	1.904	2.40	2.48
'Yb'	70	3	1.985	1.875	2.371	2.43
'Zn'	30	2	1.704	1.62	2.01	2.09
'Zr'	40	4	1.937	1.854	2.33	2.41

References Cited

Brese, N.E. and O'Keeffe, M. (1991) Bond-valence parameters for solids. Acta Crystallographica, B47, 192-197.

Hawthorne, F.C., MacDonald, D.J., Burns, P.C. (1993) Reassignment of cation site occupancies in tourmaline: Al-Mg disorder in the crystal structure of dravite. American Mineralogist, 78, 265-270.

Wright, S.E., Foley, J.A.,and Hughes, J.M. (2000) Optimization of site-occupancies in minerals using quadratic programming. American Mineralogist.

System Requirements, Program Availability

OCCQP is copyrighted software. It is available free of charge and can be obtained from this web site or directly from the authors. Version 0.39 of OCCQP consists of a single MATLAB® M-file "occqp.m" which must be run within the MATLAB computing environment in conjunction with the MATLAB Optimization Toolbox. Version 0.39 was written using MATLAB version 5.2 and works with MATLAB versions 5.2 and 5.3 on a Windows95 platform. MATLAB® and its Optimization Toolbox are commercial software products which can be purchased from The MathWorks, Inc., 24 Prime Park Way, Natick, MA 01760-1500.